Pilot Signal Design and Direct Ranging Methods for Radio Localization Using OFDM Systems

Jing, Lishuai

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Pilot Signal Design and Direct Ranging Methods for Radio Localization Using OFDM Systems

Ph.D. Dissertation

Lishuai Jing

Aalborg University
Department of Electronic Systems
Fredrik Bajers Vej 7
DK-9220 Aalborg

Thesis submitted:	January, 2015
PhD Supervisor:	Prof. Dr. Sc. Techn. Bernard H. Fleury Aalborg University
Assistant PhD Supervisor:	Assoc. Prof. Ph.D. Troels Pedersen Aalborg University
PhD Committee:	Assoc. Prof. Ph.D. Davide Dardari University of Bologna at Cesena. Research Engineer Benoit Denis CEA-Leti Minatec, Grenoble, France. Assoc. Prof. Ph.D. Troels Bundgaard Sørensen Aalborg University (Chairman).
PhD Series:	Faculty of Electronics, Aalborg University

ISBN: 978-87-7152-055-2

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2015

Abstract

Having accurate localization capability is becoming important for existing and future terrestrial wireless communication systems, in particular for orthogonal frequency-division multiplexing (OFDM) systems, such as WiMAX, wireless local area network, long-term evolution (LTE) and its extension LTE-Advanced. To obtain accurate position estimates, not only advanced estimation algorithms are needed but also the transmitted signals should be scrutinized. In this dissertation, we investigate how to design OFDM pilot signals and propose and evaluate high accuracy ranging techniques with tractable computational complexity for localization.

We first employ an important tool from radar theory, the ambiguity function, to assess the accuracy of joint delay and Doppler shift estimation using a certain pilot signal. Accordingly, an optimal pilot signal should lead to an ambiguity function with a narrow main-lobe and low side-lobes. It is found that the equispaced and equipowered pilot signal (as used in LTE) results in an ambiguity function with high side-lobes. We propose to use the Cramér-Rao bound in combination with the normalized side-lobe level (NSL) of the ambiguity function as figures of merit to devise the pilot signals. We then formulate the pilot signal design problem as a constrained optimization problem for which we propose a genetic algorithm to compute close-to-optimal solutions. The proposed method is a sound choice in a single-path scenario and a multi-path scenario with separable path components.

For scenarios where the number of path components is unknown and these components are not necessary separable, we propose a direct ranging technique using the received frequency-domain OFDM pilot signals. Compared to conventional (two-step) ranging methods, which estimate intermediate parameters such as the received signal strength, time-of-arrival, and biases introduced by non-line-of-sight (NLOS) propagation, etc., the direct ranging approach estimates the range from the received signal in one-step. This approach has the merit that it avoids LOS and first-path detection problems, the requirement of knowing the number of path components and it relaxes the separability condition of the path components. Employing a point process formulated channel model, which allows us to compute the necessary moments of the received sig-

nal, we propose and evaluate non-Bayesian and Bayesian range estimators. We show by means of Monte Carlo simulations that the proposed estimators, while exhibiting a tractable computational complexity, yield a significant ranging accuracy gain as compared to the non-coherent correlator-based estimator.

Resumé

Pålidelig lokalisering af mobile enheder bliver I stigende grad vigtig for jordbaserede trådløse kommunikationssystemer. Dette gør sig særligt gældende for såkaldte "orthogonal frequency-division multiplexing" (OFDM) systemer, såsom WiMAX, trådløse LAN netværk, LTE og udvidelsen LTE-A. For at opnå præcis estimering af position, påkræves ikke kun avancerede algoritmer, men også designet af OFDM pilot-signalerne skal undersøges. I denne afhandling undersøger vi designet af OFDM pilot signaler. Vi foreslår og evaluerer afstandsmålingsalgoritmer med høj præcision og overkommelig beregningsmæssig kompleksitet.

Vi anvender først et vigtigt værktøj fra radar teori, tvetydighedsfunktionen, til at evaluere nøjagtigheden ved samlet estimering af forsinkelse og Doppler skifte under et givet pilot signal. Et optimalt pilot signal er kendetegnet ved en tvetydighedsfunktion med smal hovedsløjfe og små sidesløjfer. Det viser sig at ved ligelig fordeling af afstand og effekt mellem pilot signalerne (som i LTE), opnås en tvetydighedsfunktion med store sidesløjfer. Vi foreslår brugen af Cramér-Rao uligheden sammen med normaliseret sidesløjfe niveau af tvetydighedsfunktionen som godhedstal for udvælgelsen af pilot signaler. Vi formulerer derefter designet af pilot signalet som et begrænset optimeringsproblem, for hvilket vi foreslår en genetisk algoritme der finder næsten-optimale løsninger. Den foreslåede metode finder anvendelse i tilfælde med kun én enkelt udbredelsesvej, samt når alle udbredelsesveje er velseparerede.

I tilfælde hvor antallet af udbredelsesveje er ukendt og udbredelsesvejene ikke nødvendigvis kan separeres, foreslår vi en metode til direkte at beregne afstande ved brug af de modtagne OFDM pilot signaler i frekvensdomænet. Konventionelle afstandsberegningsmetoder fungerer i to trin, idet de først estimerer mellemliggende parametre såsom modtaget signal styrke, ankomsttidspunkt af udbredelsesveje og bias introduceret af ikke-sigtelinje udbredelse, osv. Vores foreslåede metode fungerer, i modsætning til de konventionelle metoder, i et enkelt trin og afstanden beregnes direkte fra the modtagne signal. Fordelen ved denne fremgangsmåde er at detektion af sigtelinje signalet eller først ankomne udbredelsesvej undgås, antallet af udbredelsesveje ikke skal kendes samt krav til adskilligheden af udbredelseskomponenterne undgås. Ved

at antage en kanalmodel baseret på en punktproces foreslår og evaluerer vi Bayesianske samt ikke-Bayesianske metoder til afstandsestimering. Ved hjælp af Monte Carlo simulationer viser vi at de foreslåede metoder giver en betydelig forbedring i præcisionen af afstandsestimatet i forhold til ikke-kohærent korrelations-baserede estimering.

Preface

This thesis is submitted to the Doctoral School of Engineering and Science at Aalborg University, Denmark, in partial fulfillment of the requirements for the degree of doctor of philosophy. The first part of the thesis contains an introduction and a brief description of the contributions obtained throughout the study. As an outcome, three peer-reviewed conference papers and one journal are attached. It has been supported in part by the Faculty of Engineering and Science with PhD Stipend J. No. 562/06-8-10049 and in part by the Wireless Hybrid Enhanced Mobile Radio Estimators Phase 2 (WHERE2). It is also in part supported by the EU FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306) and in part by the VIRTUOSO cooperative research project, funded by Intel Mobile Communications, Anite, Telenor, Aalborg University, and the Danish National Advanced Technology Foundation.

I would like to express my sincere gratitude and thanks to my supervisors Professor, Dr. sc. techn. Bernard H. Fleury and Associate Professor Troels Pedersen. I am grateful for your guidance, encouragement, and advices to my PhD studies. Thank you Bernard for bringing me to the section and your precious time and effort and scientific contributions. Thanks Troels for your patience, insightful discussions, and the scientific methodologies. I would like to acknowledge my present and former colleagues and fellow PhD students at Aalborg University for many pleasant and inspiring discussions. I would like to thank our secretaries and their colleagues for the help and assistance throughout the years. I would also like to thank all my friends in China, Sweden, and Denmark for being a part of my life and all the help on the way. Last but not least, I would like to express the deepest gratefulness to my grandparents, my parents, my brothers, my sisters-in-law for their lasting support. Special thanks dedicate to my lovely wife, Heng Yang. Thanks for being there all the time.

Lishuai Jing
Aalborg University, September 6, 2015

Contents

Thesis Details

Thesis Title:	Pilot Signal Design and Direct Ranging Methods for Radio Localization Using OFDM Systems
Ph.D. Student:	Lishuai Jing
Supervisors:	Prof. Dr. Sc. Techn. Bernard H. Fleury, Assoc. Prof. Troels Pedersen, Aalborg University

The main body of this thesis consists of the following papers.

[A] Lishuai Jing, Troels Pedersen, Bernard H. Fleury, "Pilot signal design via constrained optimization with application to delay-Doppler shift estimation in OFDM systems", *51st Annual Allerton Conference on Communication, Control, and Computing*, pp. 160–166, Oct. 2013.

[B] Lishuai Jing, Troels Pedersen, Bernard H. Fleury, "Direct Ranging in Multi-path Channels Using OFDM Pilot Signals", *15th IEEE International Symposium on Signal Processing Advances in Wireless Communications*, pp. 150–154, June 2014.

[C] Lishuai Jing, Troels Pedersen, Bernard H. Fleury, "Bayesian Estimators for Direct Ranging in Multi-path Channels Using OFDM Pilot Signals" (**Submitted to**), *IEEE Trans. on Wireless Commun.*, Nov. 2014.

[D] Lishuai Jing, Troels Pedersen, Bernard H. Fleury, "Bayesian Ranging for Radio Localization with and without Line-of-Sight Detection" (**Submitted to**), *Proc. IEEE International Conference on Communications (ICC) Workshop on Advances in Network Localization and Navigation (ANLN)*, Dec, 2014.

This thesis has been submitted for assessment in partial fulfillment of the PhD degree. The thesis is based on the submitted or published scientific papers which are listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty. The thesis is not in its present form acceptable for

open publication but only in limited and closed circulation as copyright may not be ensured.

Part I

Introduction

Introduction

With soaring number of smart devices, location-based services are becoming important for terrestrial wireless systems [43]. Having location information not only becomes a demand from network operators but also is a request from the users' perspective in many scenarios. With location information of the subscribers, network operators can perform fraud detection, automated billing, radio resource management, and predict channel information etc. [29]. For users, services such as navigation and tracking, advertising, games etc. are becoming an important part of their daily lives. In addition, location information can be exploited for rescue services required from E-911 in USA and E-112 in Europe. Therefore, obtaining reliable location information can boost the users' experience that the network operators deliver and increase the success rate of rescuing.

Currently, global navigation satellite systems (GNSSs) are most often used to provide these services. Well-known systems include the American Global Positioning System (GPS), the Russian GLONASS, the European Galileo, and the Chinese COMPASS. GNSSs deliver accurate location information if at least four satellites are visible[1] to the to-be-localized terminal at the same time, see Fig. 1. In a synchronized GNSS, the geometric distance (range r) between

[1]It means that there is an unobstructed propagation path or direct path between a terminal and the satellite's antenna.

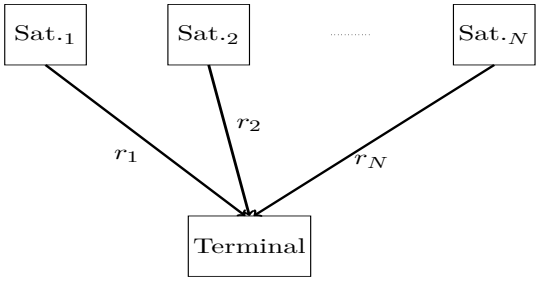


Fig. 1: To-be-localized terminal with N visible satellites (Sat.).

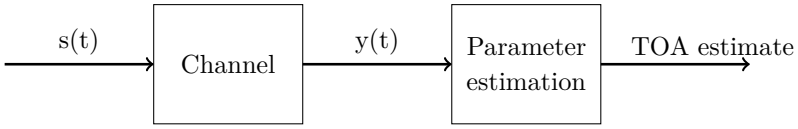


Fig. 2: A block diagram of a communication/localization system with (known) transmit signal $s(t)$, unknown channel, and received signal $y(t)$ which can be used for estimating the TOA. Note that this diagram symbols not only a satellite-based localization system but also positioning systems using other networks, for example the wireless communication networks.

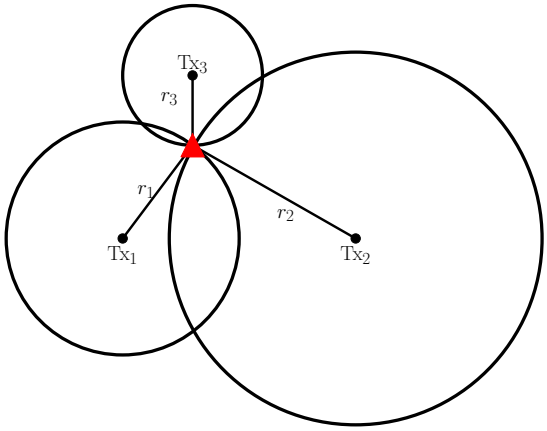


Fig. 3: An example of localization using trilateration principle on a 2D map in a noiseless scenario. The solid line indicates a constant range circle. The red triangular symbol marks the position of the to-be-localized terminal.

a satellite and the terminal is firstly inferred from the time-of-arrival (TOA). Fig. 2 shows a general system model for estimating the TOA. Typically, a known signal $s(t)$ is emitted at the terminal or a satellite. After passing through an unknown channel, it is picked up at the receiver side as $y(t)$, which can then be used to estimate the TOA. Afterwards, trilateration is used to obtain the position of a terminal based on the range estimates. Fig. 3 shows the principle of trilateration on a 2D map. From one range estimate, the terminal infers that its position is somewhere on a circle centered at a satellite position. Therefore, in an ideal case, when at least three range estimates between the terminal and the corresponding satellites are available, the intersection of the range circles provides the terminal's position unambiguously on a 2D map. To obtain a 3D coordinate, at least four satellites are required.

The GNSSs' operating condition, i.e. at least four visible satellites, is often not satisfied in indoor environments, city areas with high rise buildings, or forestry etc. In these environments, multi-path propagation and non-LOS (NLOS) condition prevail. If this happens, the positioning accuracy may be significantly degraded [44] [45].

When GNSSs fail to operate, terrestrial communication systems with localization capabilities are proposed to substitute, complement, or supplement satellite-based positioning systems [33] [45]. For example, assisted GPS (A-GPS) systems aim at improving the performance of GPS systems by exploiting the current terrestrial communication networks, such as Global System for Mobile (GSM) Communications, Wideband Code Division Multiple Access (WCDMA), and Long Term Evolution (LTE) etc. In fact, these networks have the potential to be stand-alone solutions for positioning since the deployed base stations offer a high coverage, especially in urban canyons or inside buildings where GPS precisely fails to operate.

State-of-the-art Localization Techniques

We now provide and discuss a brief overview of the state-of-the-art localization techniques. We classify them into three categories:

Fingerprinting methods employ pattern matching techniques to obtain location information [31] [45]. These methods rely on "off-line" collected and stored fingerprints, which are location-dependent parameters sampled at a specific grid of the considered environment. In the "on-line" phase, a terminal obtains its position by matching the estimated parameters with those stored in the database. Potentially, this approach is capable of providing reliable location estimation in NLOS conditions. However, it may become inaccurate when the database is outdated due to changes of the environment or movement of the transmitter or receiver.

Geometry-based algorithms exploit the received signal strength (RSS), angle of arrival (AOA), and propagation delay [45] [55] [25], etc., for ranging and positioning. Range estimators employing RSS techniques are used because of their low cost and many terminals are capable of estimating RSS. However, they require an accurate path loss model and the estimation error grows with increasing range. AOA methods reduce the number of required measuring units at the expense of demanding multiple antennas. This increases the manufacturing complexity and cost of hardware units, especially for small devices. Furthermore, when a terminal is far from the base station, small angle measurement errors may significantly degrade the positioning accuracy [45]. Localization methods based on measuring the propagation time between transmitter and receiver are widely deployed in wireless communication systems [20]. Despite requiring accurate synchronization between the transmitter and all the receivers, TOA based methods are commonly preferred for systems with large bandwidth or low complexity devices. Time-difference-of-arrival based methods release the requirement on the synchronization between the transmitter and

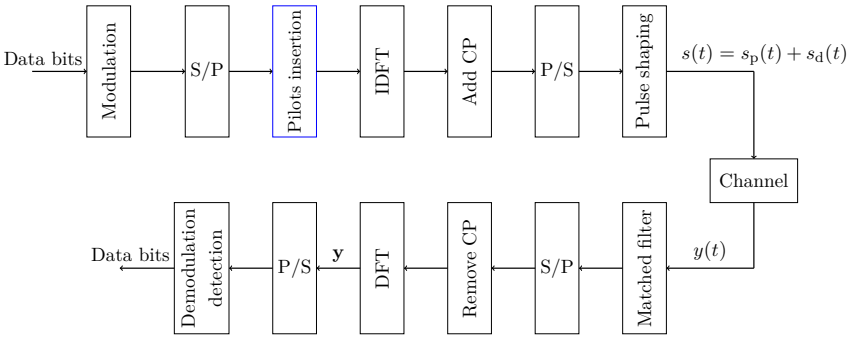


Fig. 4: Block diagram of an OFDM system base-band representation.

receiver, but they still demand accurate synchronization between all the receivers [45].

Direct positioning methods estimate a terminal position directly from the observation of the received signals [37] [36]. In these techniques, the received signals (raw measurements) from all base stations are collected and then used at a central server unit to estimate the terminal position. They potentially can improve the localization performance in multi-path environments. However, jointly processing received signals from all available base stations may introduce rather high computational complexity.

Among the above mentioned techniques, range-based localization is preferred and standardized in cellular networks [20] [25]. In this thesis, we investigate how to obtain reliable range estimates for localization in orthogonal frequency-division multiplexing (OFDM) based communication systems. We address two important aspects to achieve this goal, namely the design of optimal transmit signals (typically pilot signals) and high-accuracy ranging algorithms. In the following, we first introduce the signal model of an OFDM system to set up the estimation problem. Then we lay out the focus of this thesis.

Signal Model for OFDM Transmission

In modern communication systems, OFDM has been widely used in many existing systems, such as WLAN, LTE and its extension LTE-A [48] [33]. This technique is considered due to its flexibility in bandwidth utility, robustness to multi-path propagation, simple synchronization and equalization schemes, and its ability to achieve high data rates [1].

For simplicity, we consider a single-input single-output OFDM setup with N sub-carriers and K symbols in a frame, see Fig. 4. An OFDM symbol with time duration T is generated by multiplexing a sequence of data symbols and known pilot symbols onto N orthogonal sub-carriers with frequency spacing $\Delta f = \frac{1}{T}$.

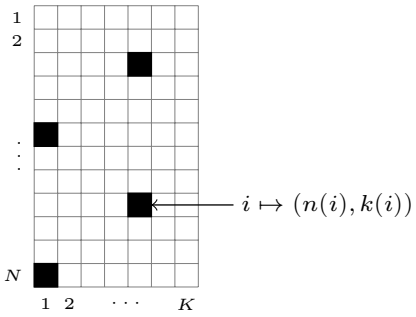


Fig. 5: The structure of an OFDM frame with $N = 12$ subcarriers and $K = 7$ symbols. Each square represents a resource element. Black and white squares indicate resource elements for pilot and data symbols, respectively. We also show the i^{th} pilot location and its mapping function.

Afterwards, time domain symbols are obtained by applying an inverse Fourier transform. Finally, a cyclic prefix of duration T_{cp} is appended to prevent inter-symbol and inter-carrier interference.

At the pilot insertion stage in Fig. 4, $N_{\text{p}} = |\mathbf{I}_{\text{p}}|$ pilots are multiplexed with data symbols to form an OFDM frame, see Fig. 5, where $\mathbf{I}_{\text{p}}$ denotes the set of pilot indices. The remaining resource elements are allocated for data symbols with indices set $\mathbf{I}_{\text{d}}$. We further define the mapping

$$\mathbf{I} \to \{1, \ldots, N\} \times \{1, \ldots, K\} : i \mapsto (n(i), k(i)),$$

where $n(i)$ and $k(i)$ specify respectively the sub-carrier and the OFDM symbol index of resource element i. Accordingly, the complex based-band signal model for the transmission of one OFDM frame reads

$$s(t) = \underbrace{\sum_{i \in \mathbf{I}_{\text{p}}} s_i(t)}_{s_{\text{p}}(t)} + \underbrace{\sum_{i \in \mathbf{I}_{\text{d}}} s_i(t)}_{s_{\text{d}}(t)} \tag{1}$$

with

$$s_i(t) = \frac{1}{\sqrt{E_{\text{p}}}} a_i e^{\jmath 2\pi n(i) \Delta f (t - k(i) T_{\text{p}})}$$

where E_{p} is the pilot signal energy, $\jmath = \sqrt{-1}$, and $T_{\text{p}} = T + T_{\text{cp}}$. In this thesis, the term pilot signal embraces the pilot pattern (i.e. placement of pilots in an OFDM frame) and signatures (i.e. amplitudes and phases).

The signal $s(t)$ is then modulated to the carrier frequency and radiated to a wireless medium, the so-called radio channel, before it is picked up at the receiver side as $y(t)$. In real environments, due to reflection, diffraction, scattering, and noise, a superposition of different replicas of the transmit signal $s(t)$

plus noise contribution is observed at the receiver side. In addition, movements of the transmitter or receiver or possibly changes in the surroundings may lead to a time-varying channel. In mathematical form, the received signal $y(t)$ is given by

$$y(t) \quad = \quad \int h(t,\tau)s(t-\tau)\mathrm{d}\tau + N(t),$$

where the noise contribution $N(t)$ is assumed to be a circular white complex Gaussian process and the time-dependent channel impulse response $h(t,\tau)$ reads

$$h(t,\tau) = q\alpha_0\delta(\tau-\tau_0)e^{-\jmath 2\pi\nu_0 t} + \sum_{l=1}^{L}\alpha_l\delta(\tau-\tau_l)e^{-\jmath 2\pi\nu_l t}, \tag{2}$$

where the lth multi-path component, $l = 0,\ldots,L$, is characterized by its complex gain α_l, delay τ_l, and Doppler shift ν_l. The LOS indicator q takes value 1 in a LOS condition and 0 otherwise. Accordingly, the channel transfer function, i.e. the Fourier transform of $h(t,\tau)$ with respect to the delay variable τ, reads

$$H(t,f) = q\alpha_0 e^{-\jmath 2\pi f\tau_0}e^{-\jmath 2\pi\nu_0 t} + \sum_{l=1}^{L}\alpha_l e^{-\jmath 2\pi f\tau_l}e^{-\jmath 2\pi\nu_l t}. \tag{3}$$

Based on these notations, we first optimize the pilot signals for estimating range-related parameters such as TOA and Doppler shift based on a single path channel assumption in Chapter 1.2. In a multi-path propagation scenario, we then propose a ranging technique, which infers range r directly from the received OFDM pilot signals in the frequency domain in Chapter 2.

Focus of the Thesis

To estimate the range for localization in OFDM communication systems, we focus on two topics, namely the design of OFDM pilot signals for estimating range related parameters and novel ranging algorithms. In the first topic, we propose close-to-optimal pilot signals for estimating delay and Doppler shift under a single path channel and a multi-path channel with separable path components. Using an important tool from radar theory, the ambiguity function, we show that the state-of-the-art pilot signals are not optimal for estimating delay and Doppler shift. In the second topic, we propose direct ranging methods, which bypass the LOS condition and first-path detection problems and estimating the number of path components.

The thesis essentially addresses two key research questions:

1. What is the optimal pilot signal $s_{\mathrm{p}}(t)$ for estimating delay and Doppler shift in single path or multi-path channels with well-separated path components?

2. Relaxing the separability condition of the path components in the channel response, how to improve ranging performance by developing novel ranging algorithms with tractable computational complexity, which overcome some challenges that the state-of-the-art range estimators are facing?

1 Pilot Signal Design for OFDM Systems

This section addresses Research Question 1. We first outline the state-of-the-art pilot signals and their selected figures of merit in OFDM based communication systems in Section 1.1. In Section 1.2, we emphasis our contributions on pilot signal design for estimating the delay and Doppler shift. In principle, the more resource elements are allocated for pilot symbols in an OFDM frame, see Fig. 5, the better the performance to be obtained is expected. But the number of allocated resource elements for data symbols is reduced that may lower the data rate. Therefore, in the optimal case, we would like to employ as few pilots as possible to keep certain constraints on a selected figure of merit satisfied. In this thesis, we employ the same number of pilots as used in a LTE system. Therefore, in the first topic of the thesis summarized in Section 1.2 and detailed in Paper A, we focus on optimizing the pilot pattern and signatures based on a proper selected channel model and figure of merit.

1.1 Pilot Signal Design for Communications

In communication systems, the traditional objective of pilot signal design is to find parsimonious pilot signals that lead to efficient channel estimation in OFDM receivers. Figures of merit such as the mean-square-error (MSE) of the channel impulse or frequency response estimates, channel capacity/throughput, outage probability, and bit error rate are widely used[2] [50]:

Information theoretic metric. By using this metric, the objective is to find optimal pilot signals that maximize Shannon's capacity. Therefore, mutual-information should be maximized with respect to the number, pattern and signature of the pilots [3] [38] [32]. Unfortunately, depending on the involved channel model, the required expression for the capacity is often difficult to obtain [50]. In some cases, however, lower bounds of the capacity can be derived. In this case, pilot signals are designed to maximize such lower bounds given the channel estimates obtained at the receiver.

Channel estimation error: MSE and Cramér-Rao bound (CRB). Accurately estimating the channel impulse or frequency response is of

[2]Depending on the application, other figures of merit exist. Here we only outline some commonly used criteria.

great importance for decoding. Therefore, pilot signals are optimized to minimize the MSE of the channel estimates [35] [8]. In addition, the CRB is also widely used, since it is desirable that the design of pilot signals does not depend on a specific estimation algorithm employed at the receiver [12]. However, if the estimation problem is non-linear, the CRB does not show at which SNR, the MSE deviates from the CRB, see Fig. 7; therefore, the CRB is only a sensible figure of merit at high SNR [18] [4].

Bit error rate (BER). For decoding purposes, the BER (or symbol error rate) is the most appropriate performance metric. It is also one of the most difficult metrics to characterize precisely. Using this metric, the optimal or close-to-optimal pilot signal minimizes the BER [9] [58].

Peak to average power ratio (PAPR). A high PAPR reduces the efficiency of power amplifiers. Accordingly, pilot signal can also be designed to minimize PAPR [19].

For single-input single-output OFDM systems, equispaced and equipowered pilot signals are shown to maximize the channel capacity [32] [38], minimize the channel estimation error [34] [35], and minimize the BER [16] [19] in their considered scenarios. The optimal pilot signal for multiple-input-multiple-output-OFDM can be found in [8] [34].

1.2 Pilot Signal Design for Delay and Doppler Shift Estimation

The demand from location-based services puts another requirement on the pilot signals: they should be additionally designed to provide accurate positioning performance. One way to improve localization performance is to obtain reliable range estimates [20]. Therefore, it may be important to estimate range-related parameters such as delay and Doppler shift. These estimates are also used for synchronization, channel prediction etc. in communication systems. In addition, delay and Doppler shift estimation is also relevant in many signal processing areas, for example sonar/radar range and speed estimation, motion detection and compensation in image processing to name a few.

We design the pilot signal $s_{\mathrm{p}}(t)$ to obtain accurate estimation of the delay and Doppler shift. For simplicity, we assume a single path channel. Accordingly, the received signal $y(t)$ reads

$$y(t) = q\alpha_0 s(t;\boldsymbol{\theta}_0) + N(t),$$

where $\boldsymbol{\theta}_0 = [\tau_0, \nu_0]$. We further assume that the receiver employs a maximum-likelihood (ML) estimator to estimate $\boldsymbol{\psi}_0 = [\boldsymbol{\theta}_0, \alpha_0]^T$ based only on the observation of the pilot signal, i.e. we set $s_{\mathrm{d}}(t) = 0$ in (1). Therefore, maximizing

the log-likelihood function of the delay and Doppler shift is equivalent to maximizing [41]

$$\left| \int s_{\mathrm{p}}^{*}(t;\boldsymbol{\theta}) y(t) \mathrm{d}t \right|^2 = \left| \underbrace{\int s_{\mathrm{p}}^{*}(t;\boldsymbol{\theta}) s_{\mathrm{p}}(t;\boldsymbol{\theta}_0) \mathrm{d}t}_{\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)} + \underbrace{\int s_{\mathrm{p}}^{*}(t;\boldsymbol{\theta}) N(t) \mathrm{d}t}_{W(\boldsymbol{\theta})} \right|^2 \tag{4}$$

in which the term $\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)$ is the so-called ambiguity function of $s_{\mathrm{p}}(t)$ [52] [18][3]. It is a two-dimensional function characterizing the similarity between the transmit signal and a version of itself shifted in delay and Doppler shift. In fact,the ambiguity function [52] is an important tool in radar theory that is used to assess the estimation accuracy in joint estimation of delay and Doppler shift of a transmit signal. Setting $\nu = 0$, we obtain the auto-correlation function of $s_{\mathrm{p}}(t)$. Selecting $\tau = 0$, we obtain the Fourier transform of the squared magnitude of the complex envelope $s_{\mathrm{p}}(t)$.

The ambiguity function has some important properties [18] [52]:

- Symmetry with respect to $\boldsymbol{\theta}_0$

$$|\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)| = |\chi(-\boldsymbol{\theta},\boldsymbol{\theta}_0)|. \tag{5}$$

- Volume invariance property

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)|^2 \mathrm{d}\tau \mathrm{d}\nu = E_{\mathrm{p}}. \tag{6}$$

- Maximum at $\boldsymbol{\theta}_0$

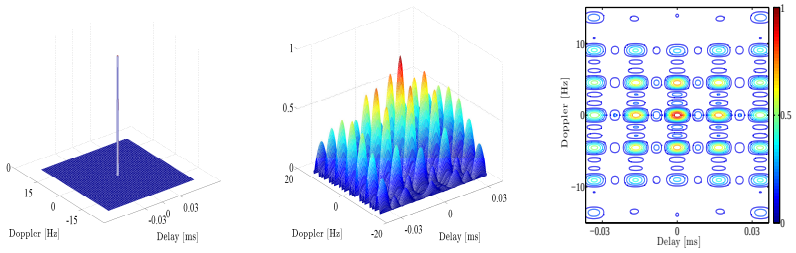
$$|\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)| \leq |\chi(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0)| = E_{\mathrm{p}}. \tag{7}$$

If $E_{\mathrm{p}} = 1$, we name $\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)$ as the normalized ambiguity function. We define the normalized side-lobe level (NSL) of the ambiguity function as the magnitude of its highest side-lobe.

Fig. 6 shows the magnitude of the ambiguity functions of a Dirac delta and a practical band-limited signal. We observe that the ambiguity functions are symmetric and exhibit their maximum at $\boldsymbol{\theta}_0$. In fact, the ambiguity function in Fig. 6(a) exhibits a desired behavior, namely a peaky main-lobe and no side-lobes. But for band-limited signals, the corresponding ambiguity functions exhibit a main-lobe and many side-lobes, see Fig. 6(b).

The ambiguity function is closely related to the achievable estimation accuracy of the delay and Doppler shift. For a nonlinear estimation problem, an example of a sketch of the typical behavior of the MSE of the delay and Doppler shift is shown in Fig. 7. The CRBs for delay and Doppler shift depend essentially on the curvature of the main-lobe of the ambiguity function at

[3]For wide-band signals, the definition of ambiguity function can be found in [30].



(a) (b) (c)

Fig. 6: Examples of the ideal ambiguity function (a) and the ambiguity function of a equipower, equispaced pilot signal (b). A contour plot of (b) is shown in (c). Without loss of generality, we set $\boldsymbol{\theta}_0 = \mathbf{0}$.

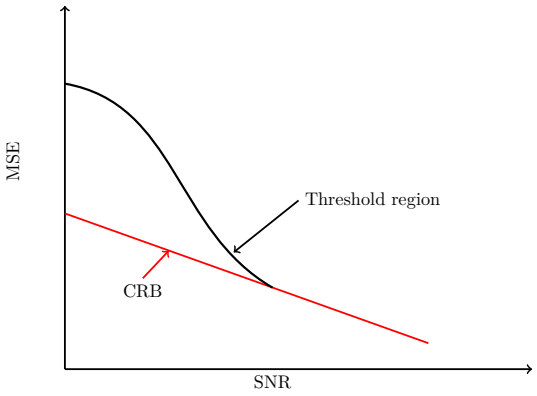


Fig. 7: MSE of the delay or Doppler shift versus SNR.

$\boldsymbol{\theta}_0$ [18]. Narrowing the main-lobe in fact reduces the CRB. Due to the volume invariance property, reducing the width of the main-lobe may increase the NSL. The NSL is closely related to the SNR threshold at which the MSE starts to significantly deviate from the CRB [4], see Fig. 7. The lower the NSL is, the more robust the estimator is towards noise. Thus, the pilot signal design problem can be formulated as: we *search for a pilot signal* $s_{\mathrm{p}}(t)$ *with an ambiguity function exhibiting a narrow main-lobe and low NSL.*

In a single path scenario, the ML estimator for delay and Doppler shift is a non-coherent correlator-based estimator, see (4). In a multi-path scenario with separable path components, if a pilot signal has an ambiguity function with sufficiently narrow main-lobe and low side-lobes, the objective function (4) exhibits L well separated dominant peaks. Therefore, the non-coherent correlator-based estimator is optimal since it achieves the CRBs [51]. Therefore, using the ambiguity function to design pilot signal is a sound approach in single-path and multi-path propagation scenarios with separable path components.

Applying the ambiguity function to design pilot signal becomes problematic in multi-path scenarios where the path components are not necessarily separable. Firstly, the computation of the CRBs for delay and Doppler shift requires the number of path components to be known. If these components are non-separable, reliably estimating its number is challenging. Even if the CRBs are computable, they only are sensible measures of the accuracy of the delay and Doppler shift at high SNR region. Secondly, it is difficult to define the NSL for a multi-dimensional ambiguity function. Accordingly, to obtain the optimal pilot signals in multi-path scenarios, we may need to resort to other tools, which is not a part of this thesis.

1.3 Contributions of the Thesis: Close-to-optimal Pilot Signals for Joint Estimation of Delay and Doppler Shift

Paper A intends to answer Research Question 1. It addresses the problem of searching for the optimal pilot signal of an OFDM system when the purpose is to estimate the delay and Doppler shift under the assumption of a single-path channel or a multi-path channel with resolvable path components. Inspired from techniques from radar theory, we propose to use the CRB and the NSL of the ambiguity function as figures of merit to devise the pilot signals. We formulate the design problem as a constrained optimization problem for which we propose a genetic algorithm that computes close-to-optimal solutions.

Several findings are reported in this paper. Firstly, the conventional equi-spaced and equipowered pilot signal, as used in LTE, is suboptimal for joint delay and Doppler estimation. It has an ambiguity function with a high NSL, which causes the ML estimator to break down at rather high SNR. Secondly, the proposed genetic algorithm generates pilot signals that minimize the NSL,

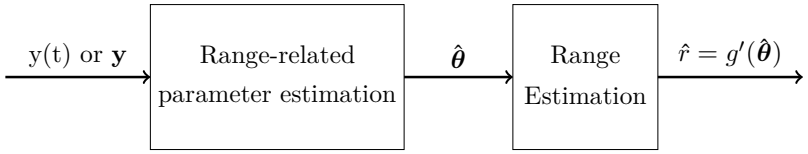


Fig. 8: Two-step ranging. In the first block, one or more range-related intermediate parameters are estimated. The second block uses these estimates to infer the range r.

while maintaining the CRBs for the delay and Doppler shift below a target value. The obtained pilot signals produce much lower NSL and CRBs as compared to the state-of-the-art pilot signals. The results clearly exemplify that the possible reduction in MSE of the delay and Doppler shift estimation can be achieved with the same number of pilots by placing them in a better manner. An important feature of the genetic algorithm is that it can generate close-to-optimal pilot signals regardless of the OFDM frame size and the number of pilots. Thirdly, the pilot pattern affects more significantly the NSL and the CRBs than the pilot signatures.

We also obtained some results on the achievable BER using the proposed pilot signals, but they are not published in open literature. We applied them to uncoded and coded (convolutional and Turbo codes) OFDM systems with winner filter applied for channel estimation. The results show that these pilot signals show inferior BER as compared to the equispaced and equipowered pilot signal. To balance the estimating accuracy of the delay and Doppler shift and the BER, constraints on the channel estimation error could potentially be added to the genetic algorithm. Another way could be to exploit more advanced channel estimation algorithms instead of the winner filter.

.

2 Direct Ranging Techniques for Localization

This section addresses Research Question 2. In wireless communication systems, path components may become non-separable due to the system bandwidth limitation. In addition, NLOS propagation prevails in indoor environments and densely populated areas in cities [56] [45] [5]. Under these conditions, how to obtain reliable ranging techniques is still an open issue.

State-of-the-art Ranging Techniques

The existing ranging techniques follow a two-step approach, see Fig. 8. Firstly, parameters, such as LOS condition, RSS, TOA, bias induced by NLOS propagation, etc., are estimated from the received signal. Then, these estimates are used for ranging. In the two-step approach, estimating TOA is essentially

equivalent to detect the first arrival path component and estimate its delay. Commonly, correlator-based, energy-detection-based [10], or multidimensional-search-based estimators [42] [49] are employed to estimate the TOA:

Correlator-based estimators are instances of the ML estimator for delay estimation under a single path channel assumption [18] [13] [33]. In multi-path scenarios, where the first path component is not necessarily the strongest or the multi-path components are not separable, the ranging accuracy of correlator-based estimators deteriorates even for ultra-wideband systems [11].

Energy-detection-based estimators offer a low-complexity method to detect the first path component [21] [2] [10]. It is difficult to select a proper threshold value as it highly depends on the noise level and the channel condition. When the threshold value is set too small, early detection may appear and when it is set too large, a miss detection of the first path may lead to large ranging errors.

Multidimensional-search-based estimators estimate all path delays and complex gains. This is similar to the channel estimation problem for communication purposes. A ML TOA method based on OFDM signals for a scenario with separable multi-path components has been proposed in [51]. Due to the assumed separability (in the delay domain) of these components, the obtained estimator converges to the correlator-based estimator. For the separability condition to hold, a large system bandwidth is needed and even in this case it is not guaranteed that all path components are separable. In [49], the authors derived the CRB on the MSE of the range estimator and investigated how OFDM signal parameters and the spacing between the multi-path components affect the bound. The CRB is a lower bound for the MSE of the TOA estimator proposed in [42], which requires estimation of the delays of all separable multi-path components. The bound in [49] and the methods in [14] [51] [42] require the knowledge of the exact number of path components, which is generally difficult to estimate reliably. Furthermore, the required multidimensional search is impractical for a realistic number of multi-path components. Potentially, lower complexity iterative schemes, such as the SAGE (space-alternating generalized expectation-maximization) algorithm [6], can be applied. These schemes, however, require the knowledge of the number of path components and might converge to a local maximum.

Under LOS conditions (with an example of power delay profile shown in Fig. 9) where the LOS component is not blocked or faded, employing state-of-the-art TOA estimators provides reliable range estimates when the SNR is sufficiently high. When the system bandwidth is insufficient, the first-path

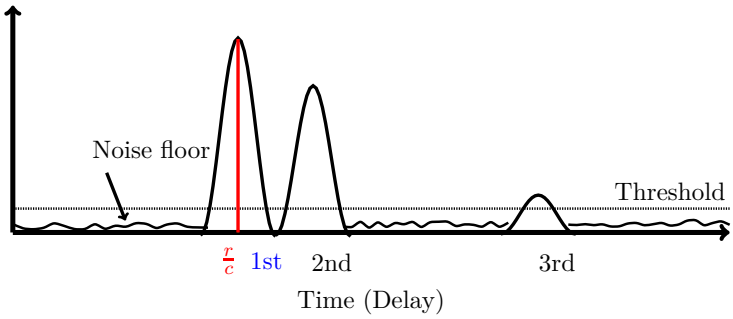


Fig. 9: An example of the power delay profile in a LOS condition. The width of the pulse depends on the system bandwidth. A threshold value is needed when energy-detection-based estimators are employed.

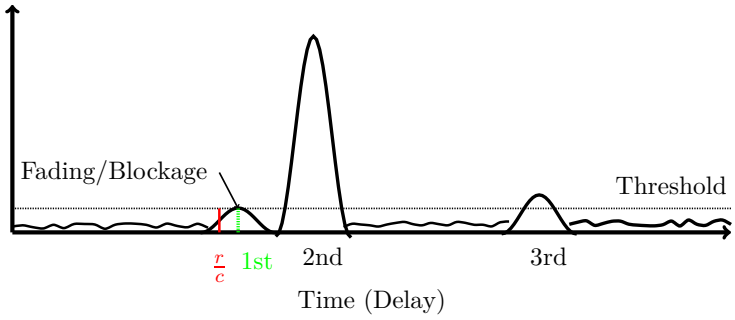


Fig. 10: An example of the power delay profile in a NLOS condition. The first path component is delayed and its amplitude is attenuated.

component is comprised of multiple non-separable path components and exhibits fading. Under these conditions, the ranging performance degrades significantly [56] [45]. In addition, the TOA estimators may be sensitive to NLOS conditions. In a NLOS condition, where the first-path is delayed as shown in Fig. 10, bias is introduced which may results in poor ranging and localization performance.

To improve the ranging accuracy using the two-step ranging methods, LOS detection and mitigation techniques are proposed to discern between the "LOS" and "NLOS" range estimates. Detailed overviews of NLOS detection and mitigation techniques can be found in [24] [45]. Such an approach is reliable provided a sufficient signal bandwidth and SNR [57] [27] [20] [46].

The LOS identification techniques can be classified into methods based on range-estimate [23], position-estimate [54], and channel statistic [24]. The channel-statistic-based methods include hypothesis-testing [22] [20] [23] and machine-learning [46] [17] approaches. These methods are brought up based

Fig. 11: Direct ranging.

on observations that certain features, such as complex amplitudes, mean delay, excess delay, and root-mean-square (RMS) delay spread differ in "LOS" or "NLOS" conditions. In [22], the kurtosis and mean excess and RMS delay of the multi-path channel are assumed to be known to form a hypothesis test to detect the LOS condition. The results in [47] indicate that employing the RMS delay spread estimate provides a robust and computationally efficient way to identify NLOS condition when ultra-wideband systems are employed. The machine-learning approach in [46] [17] employs a subset of the features of received signal amplitude and energy, rise time, mean excess delay, RMS delay spread, or kurtosis to perform NLOS identification and mitigation using support-vector-machine classifier and regressor. The NLOS identification step labels range estimates as "LOS" or "NLOS" and thereafter use this information for localization [53] [17]. The rational is that if the LOS condition can be correctly identified, this information helps improving the ranging and localization accuracy.

To provide reliable range estimates, two-step ranging methods require LOS or first-path detection, the knowledge on the number of path components or the separability condition for these components. We remark that depending on the selected range estimator, some of these requirements should be fulfilled. In communication systems with limited bandwidth and SNR, reliable detection of the LOS condition and first path and estimation of the number of path components are challenging.

Direct Ranging Techniques

To bypass the LOS and first-path detection problems and obviate the requirement on the knowledge of the number of path components and their separability condition, we propose a *direct ranging* technique as shown in Fig. 11. We assume that the Doppler shift is negligible such that the multi-path channel in (2) is time-invariant during the transmission of one OFDM symbol. Therefore, we drop the time dependency in the following representation of channel transfer function:

$$H(f) = q\alpha_0 e^{-\jmath 2\pi f \tau_0} + \sum_{l=1}^{L} \alpha_l e^{-\jmath 2\pi f \tau_l}, \tag{8}$$

where the random excess delays form a point process $\mathcal{T} = \{\tau_1, \tau_2, \ldots\}$ with intensity function $\rho(\tau)$. The shape of $\rho(\tau)$ controls the average number of

points in $\mathcal{T}$ per time unit. Depending on the specific point process assumed, the number $L = |\mathcal{T}|$ of multi-path components may be random and potentially infinite. The adoption of a channel model formulated via a point process is inspired from Turin's pioneer work [15]. The point process perspective offers not only simulation convenience to draw channel realizations, but also closed-form analytical results to characterize channel properties by using tools such as Campbell's theorem [26].

Removing the cyclic prefix and concatenating the received pilot signals in the observation vector $\mathbf{y}$, for each OFDM symbol, we obtain

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \mathbf{n}, \tag{9}$$

where $\mathbf{A} = \text{diag}\{a_1, \ldots, a_{N_\text{p}}\}$ is a diagonal matrix with a_i denoting the ith pilot symbol, $\mathbf{n}$ is a circular white complex Gaussian noise vector with component variance σ^2. The column vector $\mathbf{h}$ contains samples of the channel transfer function in (8).

Based on these notations, the direct ranging technique infers range r directly from the observation $\mathbf{y}$:

$$\hat{r} = g(\mathbf{y}), \tag{10}$$

where $g(\cdot)$ denotes a direct ranging estimator.

2.1 Maximum Likelihood Ranging

The direct ML estimator of r reads

$$\hat{r}_{ML} = \arg\max_r p(\mathbf{y}|r),$$

where $p(\mathbf{y}|r)$ denotes the likelihood function. When $p(\mathbf{y}|r)$ is unknown, an approximate ML estimator can be applied, which is of the form

$$\hat{r}_{AML} = \arg\max_r \tilde{p}(\mathbf{y}|r), \tag{11}$$

where $\tilde{p}(\mathbf{y}|r)$ denotes some approximation of the likelihood function $p(\mathbf{y}|r)$.

2.2 Bayesian Ranging Approach

If prior information on the range and LOS condition are available, we invoke Bayesian inference for ranging. We consider here the range to be a random variable with a priori pdf $p(r)$ with mean μ_r and variance σ_r^2. We model the LOS condition indicator q as a Bernoulli random variable with $p_\text{LOS} = p(q = 1)$. Therefore, Bayesian estimators such as the maximum a posteriori (MAP) estimator and minimum mean-squared error (MMSE) estimators can be applied provided that the required pdfs are available. If higher-order moments of $\mathbf{y}$ are available, a pth-order polynomial MMSE estimator [39] can be applied.

Approximate MAP Estimator

A MAP estimator selects $\hat{r}$ to maximize the posterior probability density function $p(r|\mathbf{y})$ [28]:

$$\begin{aligned}\hat{r}_{\mathrm{MAP}} &= \arg\max_r p(r|\mathbf{y}) \\ &= \arg\max_r p(\mathbf{y}|r)p(r).\end{aligned}$$

When $p(\mathbf{y}|r)$ is unknown, using an approximation $\tilde{p}(\mathbf{y}|r)$ yields

$$\hat{r}_{\mathrm{AMAP}} = \arg\max_r \ln \tilde{p}(\mathbf{y}|r) + \ln p(r). \tag{12}$$

The difference with the approximate ML estimator in (11) is that prior $p(r)$ is employed which potentially can be used to increase the ranging accuracy.

Approximate MMSE Estimator

The MMSE estimator for range r reads

$$\begin{aligned}\hat{r}_{\mathrm{MMSE}} &= E[r|\mathbf{y}] \\ &= \frac{1}{\int p(\mathbf{y}|r)p(r)\mathrm{d}r}\int rp(\mathbf{y}|r)p(r)\mathrm{d}r.\end{aligned}$$

When $p(\mathbf{y}|r)$ is unknown, we approximate $p(\mathbf{y}|r)$ with $\tilde{p}(\mathbf{y}|r)$. The resulting approximate MMSE estimator is given by

$$\hat{r}_{\mathrm{AMMSE}} = \frac{1}{\int \tilde{p}(\mathbf{y}|r)p(r)\mathrm{d}r}\int r\tilde{p}(\mathbf{y}|r)p(r)\mathrm{d}r. \tag{13}$$

For a general $p(r)$, a closed-form expression for the right-hand side of (13) may become difficult to obtain or does not exist. Instead, we may resort to numerical evaluation of the two 1D integrals in (13).

Similar to the approximate ML and MAP estimators, the performance of the approximate MMSE estimator (13) depends on the accuracy of the approximation of the likelihood function $p(\mathbf{y}|r)$. Depending on the employed pdf approximation, implementing (11) and (12) may require numerical searching procedures. Therefore, there is a tradeoff between the searching grid and the ranging accuracy. This dilemma does not exist if the approximate MMSE estimator is employed. Instead, it may require the evaluation of two 1D integrals in (13).

A pth-order Polynomial MMSE Estimator

When the pdf of the observation $\mathbf{y}$ is unknown but its higher-order moments are known or can be reliably estimated, a pth-order polynomial MMSE estimator can be employed. Such an estimator has been applied to estimate the amplitude of a real signal [39]. For complex signals, widely linear[4] MMSE ($p = 1$) and

[4]Discussions of the terminology "widely linear" can be found in [40].

widely linear quadratic MMSE ($p = 2$) estimators have been investigated in [7] and [40]. We generalize their results to complex signals. For a given $p = 1, 2, \ldots$, the pth-order polynomial MMSE estimator can be formulated as

$$\hat{r}(\mathbf{y}) = \beta_0^* + \sum_{i_1^\diamondsuit} \beta_{i_1^\diamondsuit}^* y_{i_1}^{\diamondsuit_{i_1}} + \cdots + \sum_{i_1^\diamondsuit, \ldots, i_p^\diamondsuit} \beta_{i_1^\diamondsuit, \ldots, i_p^\diamondsuit}^* y_{i_1}^{\diamondsuit_{i_1}} \cdots y_{i_p}^{\diamondsuit_{i_p}}, \tag{14}$$

where we adopt the notation $y_i^{\diamondsuit_i}$ from [40] with $\diamondsuit_i$ indicating whether or not y_i is conjugated and $i^\diamondsuit = (i, \diamondsuit_i)$. Thus, the sum over $i^\diamondsuit$ includes $2N$ terms. We recast the expression (14) in a vector form as

$$\hat{r}(\mathbf{y}) = \beta_0^* + \boldsymbol{\beta}^H \boldsymbol{z}, \tag{15}$$

where the column vector $\boldsymbol{z}$ has entries $y_{i_1}^{\diamondsuit_{i_1}} \cdots y_{i_p}^{\diamondsuit_{i_p}}$ with the associated coefficients arranged in $\boldsymbol{\beta}$.

Using the orthogonality principle to compute the coefficients in (15), we obtain

$$\hat{r}(\mathbf{y}) = \mu_r + \boldsymbol{C}_{r\boldsymbol{z}} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1} (\boldsymbol{z} - E[\boldsymbol{z}]), \tag{16}$$

where

$$\boldsymbol{C}_{r\boldsymbol{z}} = E[(r - \mu_r)(\boldsymbol{z} - E[\boldsymbol{z}])^H],$$
$$\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}} = E[(\boldsymbol{z} - E[\boldsymbol{z}])(\boldsymbol{z} - E[\boldsymbol{z}])^H].$$

As compared to the approximate MAP (12) and MMSE estimators (13), an advantage of the pth-order polynomial MMSE estimator is that no approximation on the pdf of $\mathbf{y}$ is needed. Instead, it requires that all the cross moments of $\mathbf{y}$ and r up to order p and all the moments of $\mathbf{y}$ up to order $2p$ to be known or can be reliably estimated. In addition, the covariance matrix in (16) is assumed to be invertible.

2.3 Contributions to the Thesis: Non-Bayesian and Bayesian Direct Ranging Techniques

Paper B: Maximum Likelihood Direct Ranging via Gaussian Approximations

We propose an approximate ML estimator for direct ranging. In contrast to estimators which require a multidimensional search procedure, the proposed estimator does not require the knowledge of the number of multi-path components in the channel response and these components to be separable. If the power delay spectrum of the channel and SNR are known, the computational complexity of the proposed estimator is tractable. In the single-path scenario, the non-coherent correlator-based estimator coincides with the proposed approximate ML estimator. In multi-path scenarios, the proposed estimator significantly outperforms the non-coherent correlator-based estimator.

To derive the second-order moments of the received frequency domain signal, we employ a channel model formulated via a point process approach, which allows for the application of Campbell's theorem to conveniently compute the required moments of the received signal [26]. Note that this formulation is also quite attractive for the simulation of a multi-path channel in which the path component delays and their number are random [26]. An additional finding is that both the proposed and correlator-based estimators achieve higher estimation accuracy when a random pilot pattern is employed instead of the uniform pilot pattern, as currently used in LTE. This is due to the fact that the uniform pilot signal causes high side-lobes in the objective function. This might be eliminated by selecting a proper search range for the objective function, which, however, would lead to significant constraints on the obtainable estimation range.

Paper C: Bayesian Ranging with Known LOS Condition

We employ prior information of range r for ranging assuming the LOS condition is known. In contrast to the ML estimator, we address the direct ranging problem via Bayesian estimators, namely MAP, linear MMSE, and MMSE estimators. The first finding of this work is that these estimators cannot be directly applied since the pdf of the received frequency-domain OFDM signal is unknown. Instead, we propose approximate versions of these estimators which employ approximations on the pdf of received pilot signals in the frequency domain.

Employing any multi-path channel model which makes the computation of the first- and second-order moments of the received signal possible, approximate MAP and MMSE estimators can be applied. If the channel model is formulated as a Poisson point process (the classical Turin's channel model), we compute all moments of the received signal using Campbell's theorem. This is a remarkable benefit which allows us to apply the pth-order polynomial MMSE estimator. We find that the standard widely linear MMSE estimator ($p = 1$) is inapplicable. Instead, we apply the widely linear quadratic MMSE estimator ($p = 2$), which is unbiased according to the orthogonality principle. Secondly, all the proposed estimators have merits, such as they do not rely on a first-path detection, any separability condition, and the knowledge of the number of multi-path components. Thirdly, simulation results show that the proposed estimators are fairly robust against channel model mismatches.

Paper D: Bayesian Ranging with and without LOS Detection

We propose ranging methods with and without LOS detection. In contrast to Paper C, we model the LOS condition parameter as a Bernoulli random variable instead of a known value. We propose approximate MAP and MMSE

estimators which are derived by approximating the pdf of the received signal. Through Monte Carlo simulations, we observe that the approximate MMSE estimators outperform the approximate MAP estimators in terms of RMSE. Furthermore, including LOS detection in the estimators, while adding computational complexities, has no major impact on the ranging performance at least for the employed pdf approximations. Our simulation study indicates that there is a potential for improving the ranging performance by employing better pdf approximations.

Summary of Main Findings, Conclusions, and Outlook

When the ambiguity function is used to facilitate designing pilot signals for delay and Doppler estimation, pilot signals resulting in narrow main-lobe and low NSL are required. State-of-the-art pilot signals, such as the equispaced and equipowered pilot signals, are suboptimal since their ambiguity functions exhibit high NSLs. Using the CRB and the NSL of the ambiguity function as figures of merit, the proposed genetic algorithm yields pilot signals that exhibit simultaneously low CRBs for delay and Doppler shift and low NSL. The algorithm computes close-to-optimal pilot signals offline regardless of the size of the OFDM frame and the number of pilots. Since these signals have ambiguity functions with much lower NSL, employing them increases the robustness of the non-coherent correlator-based estimator towards noise.

Using the ambiguity function as a tool to design pilot signal is a sound approach in single-path and multi-path scenarios with separable path components. Under these conditions, the ML estimator for delay and Doppler shift estimation is the non-coherent correlator-based estimator. If the system bandwidth is not high enough to resolve the multi-path components, employing an non-coherent correlator-based estimator leads to poor ranging performance.

To cope with this scenario, we propose direct ranging methods that bypass the LOS condition and first-path detection problems and obviate the requirement of knowing the number of path components in the channel response and any separability condition on these components. If the first- and second-order moments of the received signal are available, approximate ML estimators can be applied for direct ranging. When prior information on the range is available, approximate MAP and MMSE estimators are applied. If we further have access to up to $2p$th-order moments of the received signal, the unbiased pth-order polynomial MMSE estimator can be employed. Using Turin's classical channel model, all moments of the received signal can be computed analytically via Campbell's theorem. In addition, the RMSE of the polynomial MMSE estimator can be obtained in a closed-form expression. This is not the case for the

approximate ML, MAP, and MMSE estimators.

Furthermore, we propose approximate MAP and MMSE estimators that inherently take the prior of the LOS into account and thereby bypass the LOS detection problem. Monte Carlo simulations show that this approach and the approach employing a LOS detector yield similar ranging accuracies considering the employed pdf approximations. But a ranging method without LOS detection is advantageous in computational complexity.

In short, we show by means of Monte Carlo simulations that 1) applying the proposed pilot signals increases the robustness of the correlator-based estimators towards noise; 2) the proposed direct ranging estimators significantly outperform the non-coherent correlator-based estimator in a multi-path channel. These findings trigger the forthcoming question: what is the ranging accuracy using the proposed pilot signals and estimators if they are applied to measurements obtained from practical scenarios? To validate the performance of these estimators, a measurement campaign could be conducted. It would also be of high relevance to show how much the improved ranging accuracy (if there is any) translated to a superior localization accuracy is.

The performance of the proposed range estimators is inherently limited by the underlying pdf approximations. When a multi-path channel with low average number of path components is encountered, the Gaussian approximation of the conditional pdf of the received signal is poor. In this or similar cases, a straightforward way to improve the performance of the proposed estimators is to investigate more accurate approximations that take into account higher-order moments of the received signal.

In this thesis, we assume that the SNR and the delay power spectrum are known to apply the direct ranging estimators. In future contributions, one may include the estimation of the parameters of the delay power spectrum. The model mismatch analysis in Papers B and C has shown that estimating the parameters up to a certain accuracy causes negligible performance loss. Therefore, having unreliable estimates of these parameters is not critical to the performance of the proposed estimators. Another aspect not considered in this contribution that can be the subject of future investigations is the distribution of range errors. This distribution is widely used to derive range-based localization algorithms. As with other ranging methods, the exact distribution of the range errors of the proposed estimators is unknown. Thus, one must resort to empirical models of the pdf of the range errors. Since direct ranging methods do not rely on LOS and first-path detection, potential range errors resulting from missed LOS and first-path detection, which are often considered in error models, are avoided. This rationale leads us to conjecture that direct ranging techniques give rise to differently distributed range errors compared to existing ranging methods.

The proposed methods for designing pilot signals and estimating the range are not only relevant to improving ranging accuracy using wireless communi-

cation networks. They can also be applied to improve synchronization performance in communication or localization systems. In challenging scenarios where multi-path propagation prevails and LOS condition may not appear, we may also apply the proposed methods to increase the estimation accuracy of the target location and velocity for radar/sonar applications. In addition, the proposed methods potentially can be used to improve accuracy of the delay estimation for speech signal processing.

References

[1] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[2] A. Rabbachin, I. Oppermann and B. Denis. ML Time-of-Arrival estimation based on low complexity UWB energy detection. In *IEEE International Conference on Ultra-Wideband*, pages 599–604, Sept. 2006.

[3] S. Adireddy, L. Tong, and H. Viswanathan. Optimal placement of training for frequency-selective block-fading channels. *IEEE Trans. Inf. Theory*, 48(8):2338–2353, Aug. 2002.

[4] F. Athley. Threshold region performance of maximum likelihood direction of arrival estimators. *IEEE Trans. Signal Process.*, 53(4):1359–1373, Apr. 2005.

[5] B. Denis, J. Keignart, and N. Daniele. Impact of NLOS propagation upon ranging precision in UWB systems. In *IEEE Conference on Ultra Wideband Systems and Technologies*, pages 379–383, Nov. 2003.

[6] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen. Channel parameter estimation in mobile radio environments using the SAGE algorithm. *IEEE J. Sel. Areas Commun.*, 17(3):434–450, Mar. 1999.

[7] B. Picinbono and P. Chevalier. Widely linear estimation with complex data. *IEEE Trans. Signal Process.*, 43(8):2030–2033, Aug. 1995.

[8] I. Barhumi, G. Leus, and M. Moonen. Optimal training design for MIMO OFDM systems in mobile wireless channels. *IEEE Trans. Signal Process.*, 51(6):1615–1624, June 2003.

[9] X. Cai and G. Giannakis. Error probability minimizing pilots for OFDM with M-PSK modulation over Rayleigh-fading channels. *IEEE Trans. Veh. Technol.*, 53(1):146–155, Jan. 2004.

[10] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win. Ranging with ultrawide bandwidth signals in multi-path environments. *Proc. IEEE*, 97(2):404–426, 2009.

[11] D. Dardari and M. Z. Win. Threshold-based time-of-arrival estimators in UWB dense multipath channels. In *IEEE International Conference on Communications (ICC)*, volume 10, pages 4723–4728, June 2006.

[12] M. Dong and L. Tong. Optimal design and placement of pilot symbols for channel estimation. *IEEE Trans. Signal Process.*, 50(12):3055–3069, Dec. 2002.

[13] Donglin Wang and Fattouche, M. OFDM Transmission for Time-Based Range Estimation. *IEEE Signal Process. Lett.*, 17(6):571–574, 2010.

[14] F. Zhao, W. Yao, C. C. Logothetis, and Y. Song. Super-resolution TOA Estimation in OFDM Systems for Indoor Environments. In *IEEE International Conf. on Networking, Sensing and Control*, Apr. 2007.

[15] G. L. Turin, F. D. Clapp, T. L. Johnston, S. B. Fine, and D. Lavry. A statistical model of urban multipath propagation. *IEEE Trans. Veh. Technol.*, 21(1):1–9, 1972.

[16] M.-G. Garcia, S. Zazo, and J. Paez-Borrallo. Pilot patterns for channel estimation in OFDM. *Electronics Letters*, 36(12):1049–1050, jun 2000.

[17] H. Wymeersch, S. Marano, W. Gifford, and M. Z. Win. A machine learning approach to ranging error mitigation for UWB localization. *IEEE Trans. Commun.*, 60(6):1719–1728, 2012.

[18] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory - Part III.* John Wiley & Sons, 2001.

[19] S. Hosokawa, S. Ohno, K. Teo, and T. Hinamoto. Pilot tone design for peak-to-average power ratio reduction in OFDM. In *IEEE International Symposium on Circuits and Systems*, pages 6014 – 6017 Vol. 6, May 2005.

[20] I. Guvenc and Chia-Chin Chong. A survey on TOA based wireless localization and NLOS mitigation techniques. *IEEE Communications Surveys Tutorials*, 11(3):107–124, Aug. 2009.

[21] I. Guvenc and Z. Sahinoglu. Threshold selection for UWB TOA estimation based on kurtosis analysis. *IEEE Communications Letters*, 9(12):1025–1027, Dec. 2005.

[22] I. Guvenc, Chia-Chin Chong, and F. Watanabe. NLOS identification and mitigation for UWB localization systems. In *IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1571–1576, Mar. 2007.

[23] J. Borras, P. Hatrack, and N. Mandayam. Decision theoretic framework for NLOS identification. In *Proc. IEEE Int. Technol. Conf.*, volume 2, pages 1583–1587 vol.2, May 1998.

[24] J. Khodjaev, Y. Park, and A. Saeed Malik. Survey of NLOS identification and error mitigation problems in UWB-based positioning algorithms for dense environments. *Annals of telecommunications*, 65(5-6):301–311, 2010.

[25] J. Yan, C. Tiberius, G. Janssen, P. Teunissen, and G. Bellusci. Review of range-based positioning algorithms. *IEEE Aerosp. Electron. Syst. Mag.*, 28(8):2–27, Aug. 2013.

[26] M. L. Jakobsen. *Modeling and Analysis of Stochastic Radio Channels*. PhD thesis, Aalborg University, 2014.

[27] Junyang Shen and A. F. Molisch. Indirect path detection based on wireless propagation measurements. *IEEE Trans. Wireless Commun.*, 11(12):4482–4493, Dec. 2012.

[28] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall PTR, 1998.

[29] L. Brunel, M. Plainchault, N. Gresset, A. Dammann, C. Mensing, and R. Raulefs. Inter-cell interference coordination and synchronization based on location information. In *7th Workshop on Positioning Navigation and Communication (WPNC)*, pages 224–232, Mar. 2010.

[30] L. Weiss. Wavelets and wideband correlation processing. *IEEE Signal Process. Mag.*, 11(1):13–32, Jan. 1994.

[31] M. Triki and D. Slock. Mobile Localization for NLOS Propagation. In *Proc. IEEE 18th Int. Symposium on Personal, Indoor and Mobile Radio Commun., PIMRC*, pages 1–4, Sept. 2007.

[32] X. Ma, G. Giannakis, and S. Ohno. Optimal training for block transmissions over doubly selective wireless fading channels. *IEEE Trans. Signal Process.*, 51(5):1351–1366, May 2003.

[33] C. Mensing. *Location Determination in OFDM Based Mobile Radio Systems*. PhD thesis, Technische Universität München, 2013.

[34] H. Minn, N. Al-Dhahir, and Y. Li. Optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and phase noise. *IEEE Trans. Commun.*, 54(10), Oct. 2006.

[35] R. Negi and J. Cioffi. Pilot tone selection for channel estimation in a mobile OFDM system. *IEEE Trans. Consum. Electron.*, 44(3):1122–1128, Aug. 1998.

[36] O. Bar-Shalom and A. Weiss. Direct position determination of OFDM signals. In *IEEE 8th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)* , pages 1–5, June 2007.

[37] O. Bialer, D. Raphaeli, and A. J. Weiss. Maximum-likelihood direct position estimation in dense multipath. *IEEE Trans. Veh. Technol.*, 62(5):2069–2079, June 2013.

[38] S. Ohno and G. Giannakis. Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels. *IEEE Trans. Inf. Theory*, 50(9), Sept. 2004.

[39] P. Bondon, M. Benidir, and B. Picinbono. A nonlinear approach to estimate the amplitude of a signal. In *IEEE International Conf. on Acoustics, Speech, and Signal Process.*, volume 5, pages 301–304, Mar. 1992.

[40] P. Schreier and L. Scharf. *Statistical Signal Processing of Complex-Valued Data*. Cambridge University Press, 2010.

[41] H. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1994.

[42] R. Adam and P. A. Hoeher. Semi-blind channel estimation for joint communication and positioning. *10th Workshop on Positioning Navigation and Commun.*, pages 1–5, 2013.

[43] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch. Location-Aware Communications for 5G Networks: How location information can improve scalability, latency, and robustness of 5G. *IEEE Signal Process. Mag.*, 31(6):102–112, Nov. 2014.

[44] R. Ercek, P. De Doncker, and F. Grenez. NLOS-multipath effects on pseudo-range estimation in urban canyons for GNSS applications. In *First European Conference on Antennas and Propagation*, pages 1–6, Nov. 2006.

[45] R. Zekavat and R. M. Buehrer. *Handbook of Position Location: Theory, Practice and Advances*. IEEE Series on Digital & Mobile Communication. Wiley, 2011.

[46] S. Marano, W. Gifford, H. Wymeersch, and M. Win. NLOS identification and mitigation for localization based on UWB experimental data. *IEEE J. Sel. Areas Commun.*, 28(7):1026–1035, 2010.

[47] S. Venkatesh and R. Buehrer. Non-line-of-sight identification in ultra-wideband systems based on received signal statistics. *IET Microwaves, Antennas Propagation*, 1(6):1120–1130, Dec. 2007.

[48] A. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location. *IEEE Signal Process. Mag.*, 22(4):24–40, 2005.

[49] T. Wang, Y. Shen, S. Mazuelas, H. Shin and M. Z. Win. On OFDM ranging accuracy in multipath channels. *IEEE Systems Journal*, PP(99):1–11, 2013.

[50] L. Tong, B. Sadler, and M. Dong. Pilot-assisted wireless transmissions: general model, design criteria, and signal processing. *IEEE Signal Process. Mag.*, 21(6):12–25, Nov. 2004.

[51] D. Wang, M. Fattouche, and F. Ghannouchi. Fundamental limit of OFDM range estimation in a separable multipath environment. *Circuits, Systems, and Signal Processing*, 31(3), 2012.

[52] P. Woodward. *Probability and Information Theory, with Applications to Radar.* Pergamon Press, 1953.

[53] Y. Qi, K. Hisashi, and H. Suda. On time-of-arrival positioning in a multipath environment. *IEEE Trans. Veh. Technol.*, 55(5):1516–1526, 2006.

[54] Yung-Hoon Jo, Joon-Yong Lee, Dong-Heon Ha, and Shin-Hoo Kang. Accuracy enhancement for UWB indoor positioning using ray tracing. In *IEEE/ION Position, Location, And Navigation Symposium*, pages 565–568, Apr. 2006.

[55] Z. Sahinoglu, S. Gezici, and I. Güvenc . *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols.* Cambridge University Press, 2008.

[56] Z. Sahinoglu, S. Gezici, and I. Güvenc. *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols.* Cambridge University Press, 2008.

[57] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik. Non-line-of-sight identification and mitigation using received signal strength. *IEEE Trans. Wireless Commun.*, (99):1–1, Nov. 2014.

[58] W. Zhang, X.-G. Xia, and P. Ching. Optimal training and pilot pattern design for OFDM systems in rayleigh fading. *IEEE Trans. Broadcast.*, 52(4):505–514, Dec. 2006.

Part II

Papers

Paper A

Pilot Signal Design via Constrained Optimization with Application to Delay-Doppler Shift Estimation in OFDM Systems

Lishuai Jing, Troels Pedersen, Bernard H. Fleury

The paper has been published in the
51st Annual Allerton Conference on Communication, Control, and Computing
Oct. 2 - Oct. 4, Monticello, IL, USA, 2013.

The layout has been revised.

Abstract

We address the problem of searching for the optimal pilot signal, i.e. pattern and signature, of an orthogonal frequency-division multiplexing (OFDM) system when the purpose is to estimate the delay and Doppler shift under the assumption of a single-path propagation channel. This problem is relevant for synchronization and for time-based localization using said signals. We propose to use the Cramér-Rao bound and the normalized side-lobe level (NSL) of the ambiguity function as figures of merit to devise the pilot signals. We formulate the design problem as a constrained optimization problem for which we propose a genetic algorithm that computes close-to-optimal solutions. Simulation results demonstrate that the proposed algorithm can efficiently find pilot signals that outperform the state-of-the-art pilot signals in both single-path and multi-path propagation scenarios. In addition, we demonstrate that data interference causes a performance loss if a standard non-coherent correlator is used. The results also indicate that the pilot pattern impacts the estimator's performance more than the pilot signature.

1 Introduction

In Orthogonal Frequency-Division Multiplexing (OFDM) systems, data signals are embedded in an OFDM frame together with pilot signals which are used to acquire channel information [12]. In this contribution, the term pilot signal embraces the pilot pattern (i.e. the placement of pilots in the time-frequency grid) and the pilot signature (i.e. pilot amplitudes and phases). The traditional objective of pilot signal design is to find parsimonious pilot signals that lead to efficient channel estimation in OFDM receivers. A comprehensive survey of pilot signal design can be found in [12]. Equispaced and equipowered pilot signals are shown to maximize the channel capacity, minimize the channel estimation error, and minimize the bit error rate for the considered scenarios, see [12] and references therein.

The last ten years have witnessed a steady increasing endeavor in research on localization using terrestrial wireless systems, especially long-term evolution (LTE) and its extension LTE-A. The deployment of localization capabilities in terrestrial wireless systems is aimed at substituting, complementing, or supplementing satellite-based positioning systems in scenarios where the latter systems are unable to operate [10] [3] [5]. These localization features put additional requirements on the pilot signals transmitted by these wireless systems: pilot signals should additionally be designed to optimize positioning capabilities. Position-bearing channel parameters commonly exploited for localization are the received signal strength (RSS), the propagation delay, and the angle of arrival (AOA) [10] [5]. Time-of-arrival (TOA) and time-difference-of-arrival

(TDOA) based positioning methods rely on estimates of the propagation delay between the reference stations and the mobile station to be localized. Doppler shift estimate can be used to extract the relative velocity for navigation and thereby to enhance the positioning accuracy [3]. In this contribution, we focus on pilot assisted delay and Doppler shift estimation in OFDM for the purpose of synchronization and localization.

In radar theory, the ambiguity function [7] of the transmit signal is an important tool for assessing the accuracy of the joint estimation of the delay and Doppler shift. To achieve good estimation accuracy, it is mandatory that the ambiguity function exhibits a narrow main-lobe and low side-lobes. However, these two features are contradictory due to the volume invariance property [7]. The ambiguity function of equispaced and equipowered pilot signals does not fulfill the second of these requirements: it exhibits high side-lobes (see Fig. A.2 in Section 5). Two approaches have been proposed in the literature to obtain pilot signals with a "good" ambiguity function in the aforementioned sense. The first approach consists in using pilot patterns that belong to the class of "perfect periodic" Costas arrays [6]. This class is an extension of the class of Costas arrays. Costas arrays leads to an ambiguity function with low side-lobes away from the main-lobe, though high side-lobes remain near the main-lobe [4]. A limitation of the Costas arrays is their inherent constraint: the array must be square and the number of pilots must equal the array length. The class of "perfect periodic" Costas arrays [6] allow for alleviating this constraint. The second approach, proposed in [13], is to use a genetic algorithm to design pilot signals for one OFDM symbol that yields an autocorrelation function—the delay ambiguity function in our terminology—with low side-lobes. The objective function that the algorithm attempts to minimize is a linear combination of the maximum side-lobe magnitude and the 3 dB main-lobe width of the delay ambiguity function.

Inspired by the above two approaches, we consider in this contribution the constrained optimization problem of designing pilot signals that yield a delay-Doppler ambiguity function with low side-lobes, while keeping the Cramér-Rao bounds (CRBs) for the estimation of the delay and Doppler shift below a prescribed level. We propose a genetic algorithm to compute close-to-optimal solutions. For a given number of pilots, the algorithm can efficiently find pilot signals which yield lower side-lobes and CRBs than the corresponding values achieved with equispaced, equipowered pilot signals and "perfect periodic" Costas arrays. We provide simulation results showing that the pilot signals designed with the genetic algorithm lead to a better estimation accuracy compared to the accuracy achieved by using "perfect periodic" Costas arrays in both single-path and multipath channels when the delay-Doppler estimator is implemented via a standard (pilot-based) correlator. The results also show that the pilot pattern affects the estimator performance more than the pilot signature and that the data symbols affect the threshold region performance of

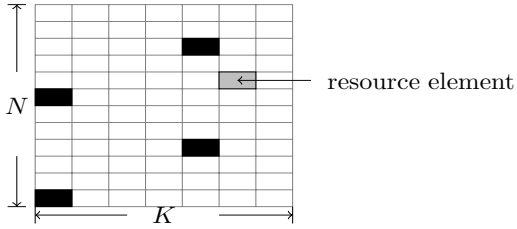


Fig. A.1: The structure of an OFDM frame with $N = 12$ subcarriers and $K = 7$ OFDM symbols. One box stands for one resource element. Black boxes indicate pilot symbols and white boxes indicate data symbols.

the correlator-based delay-Doppler estimator.

2 Signal Model

We consider a single-input single-output OFDM setup with N subcarriers and K symbols in a frame as the example shown in Fig. A.1. An OFDM symbol with time duration T is generated by multiplexing a sequence of data symbols and known pilot symbols onto N orthogonal sub-carriers. Afterward, the time domain symbols are obtained by using an inverse Fourier transform. Finally, a cyclic prefix of duration T_{cp} is appended to prevent inter-symbol and inter-carrier interference. The total duration of an OFDM symbol is thus $T_{\mathrm{p}} = T + T_{\mathrm{cp}}$. The adjacent sub-carrier spacing is $\Delta f = \frac{1}{T}$.

An OFDM frame consists of a total of NK so-called resource elements indexed by the set $\mathbf{I} = \{1, 2, \ldots, NK\}$. Of these resource elements, $N_{\mathrm{p}} = |\mathbf{I}_{\mathrm{p}}|$ are pilots indexed by $\mathbf{I}_{\mathrm{p}}$ and $N_{\mathrm{d}} = |\mathbf{I}_{\mathrm{d}}|$ are allocated to data indexed by $\mathbf{I}_{\mathrm{d}}$. We further define the mapping

$$\mathbf{I} \to \{1, \ldots, N\} \times \{1, \ldots, K\} : i \mapsto (n(i), k(i)), \tag{A.1}$$

where $n(i)$ and $k(i)$ specify the subcarrier and the OFDM symbol respectively of resource element i. The OFDM signal reads in complex baseband notation:

$$\begin{aligned} s(t) &= s_{\mathrm{p}}(t) + s_{\mathrm{d}}(t) \\ &= \sum_{i \in \mathbf{I}_{\mathrm{p}}} s_i(t) + \sum_{i \in \mathbf{I}_{\mathrm{d}}} s_i(t) \end{aligned} \tag{A.2}$$

with $s_i(t) = a_i e^{\jmath 2\pi n(i) \Delta f (t - k(i) T_{\mathrm{p}})} 1(\frac{t}{T_{\mathrm{p}}} - k(i) \in [-\frac{1}{2}, \frac{1}{2}])$. Here, a_i is the ith transmit symbol, $\jmath = \sqrt{-1}$, and $1(\cdot)$ denotes the indicator function.

Assuming transmission across a multipath propagation channel, the received signal reads

$$Y(t) = \sum_{l=0}^{L-1} \alpha_l s(t, \boldsymbol{\theta}_l) + N(t) \tag{A.3}$$

with $s(t, \boldsymbol{\theta}_l) = s(t-\tau)e^{\jmath 2\pi\nu_l t}$ where $s(t-\tau) = s_{\mathrm{p}}(t-\tau_l) + s_{\mathrm{d}}(t-\tau_l)$. The lth multipath component is characterized by its complex weight α_l, delay τ_l, and Doppler shift ν_l. We concatenate the later two parameters in the vector $\boldsymbol{\theta}_l = [\tau_l, \nu_l]^T$. The noise contribution $N(t)$ is assumed to be a circular white complex Gaussian process with autocorrelation

$$\mathrm{E}[N(t)N^*(t+\tau)] = N_0\delta(\tau), \tag{A.4}$$

where $\mathrm{E}[\cdot]$ denotes expectation, $(\cdot)^*$ stands for complex conjugation, N_0 is a positive constant, and $\delta(\cdot)$ is the Dirac delta function.

3 Maximum-Likelihood Estimation of Delay and Doppler Shift

In this section, we first derive the joint maximum likelihood estimator of the delay and Doppler shift in an OFDM scenario with pilot-only transmission across a single-path propagation channel. Then, we define the pilot ambiguity function and derive the CRBs for the estimation of the delay and Doppler shift. Finally, we propose a constrained optimization problem for pilot signal design.

We assume a single-path propagation channel ($L = 1$) with complex gain α_0, delay τ_0 and Doppler shift ν_0[1]. Furthermore, the OFDM frame duration is short enough so that α_0, τ_0 and ν_0 are constant during one OFDM frame. Under these assumptions, the channel time-frequency response is flat in frequency, but varies from one OFDM symbol to another due to the Doppler shift. We define the signal to noise ratio (SNR) $\gamma = \frac{E_{\mathrm{p}}}{N_0}$ with $E_{\mathrm{p}} = \int |s_{\mathrm{p}}(t)|^2 \mathrm{d}t$.

We further assume that the receiver estimates the unknown parameter vector $\boldsymbol{\psi}_0 = [\boldsymbol{\theta}_0, \alpha_0]^T$ based only on the observation of the pilot signal, i.e. we set $s_{\mathrm{d}}(t) = 0$, and thus $s(t) = s_{\mathrm{p}}(t)$ and $s(t; \boldsymbol{\theta}_0) = s_{\mathrm{p}}(t; \boldsymbol{\theta}_0)$. From (A.3) and with the above assumptions, the log-likelihood function of $\boldsymbol{\psi} = [\boldsymbol{\theta}, \alpha]^T$ reads [9]

$$\tilde{\Lambda}(\boldsymbol{\psi}; Y(t)) = \frac{2}{N_0}\mathcal{R}\left\{\alpha^*\Lambda(\boldsymbol{\theta}; Y(t))\right\} - \frac{|\alpha|^2}{N_0}\int |s_{\mathrm{p}}(t; \boldsymbol{\theta})|^2 \mathrm{d}t. \tag{A.5}$$

In this expression, $\mathcal{R}\{\cdot\}$ and $|\cdot|$ denote respectively the real part and the absolute value of the argument and $\Lambda(\boldsymbol{\theta}; Y(t)) = \int s_{\mathrm{p}}^*(t; \boldsymbol{\theta}) Y(t) \mathrm{d}t$. Note that the term $\int |s_{\mathrm{p}}(t; \boldsymbol{\theta})|^2 \mathrm{d}t = E_{\mathrm{p}}$ that arises in the log-likelihood function does not depend on $\boldsymbol{\theta}$. Given the pilot signal observation $Y(t) = y(t)$, the joint maximum likelihood (ML) estimator of the delay, Doppler and complex gain is

$$\hat{\boldsymbol{\psi}}_0 = \arg\max_{\boldsymbol{\psi}} \tilde{\Lambda}(\boldsymbol{\psi}; y(t)). \tag{A.6}$$

[1] We will return to the multipath scenario in Section 5.

The estimation problem in (A.6) is separable:

$$\hat{\boldsymbol{\theta}}_0 = \arg\max_{\boldsymbol{\theta}} |\Lambda(\boldsymbol{\theta}; y(t))|^2 \tag{A.7}$$

$$\hat{\alpha} = \frac{\Lambda(\hat{\boldsymbol{\theta}}_0; y(t))}{E_\mathrm{p}}. \tag{A.8}$$

Therefore, to estimate the delay and Doppler shift, we need to compute $\Lambda(\boldsymbol{\theta}; y(t))$. In practice, this computation can be implemented via a correlator which correlates the observed signal $y(t)$ with the delayed and Doppler shifted replicas of the pilot signal.

3.1 Ambiguity Function of Pilot Signals

We can rewrite the objective function in (A.7) as

$$\left| \underbrace{\int s_\mathrm{p}^*(t;\boldsymbol{\theta}) s_\mathrm{p}(t;\boldsymbol{\theta}_0)\mathrm{d}t}_{\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0)} + \underbrace{\int s_\mathrm{p}^*(t;\boldsymbol{\theta}) N(t)\mathrm{d}t}_{W(\boldsymbol{\theta})} \right|^2. \tag{A.9}$$

The term $\chi(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ is the so-called ambiguity function of $s_\mathrm{p}(t)$ [14] and $W(\boldsymbol{\theta})$ is a zero mean colored Gaussian process.

The ambiguity function of the pilot signal limits the accuracy of the estimation of $\boldsymbol{\theta}_0$. To minimize the estimation error, $s_\mathrm{p}(t)$ shall be designed such that its ambiguity function exhibits a narrow main-lobe centered at $\boldsymbol{\theta}_0$ and low side-lobes [14]. However, due to the volume invariance property ($\iint |\chi(\boldsymbol{\theta}, \boldsymbol{\theta}_0)|^2 \mathrm{d}\tau \mathrm{d}\nu = 1$), the design involves a trade-off between the width of the main-lobe and the magnitude of the side-lobes [7]. Thus if $s_\mathrm{p}(t)$ is selected such that its ambiguity function exhibits a narrow main-lobe, high side-lobes may appear and vice-versa.

3.2 Fisher Information and Cramér-Rao Bound

In the subsequent investigation, we consider the real vector $\tilde{\boldsymbol{\psi}} = [\boldsymbol{\theta}, \mathcal{R}(\alpha), \mathcal{I}(\alpha)]^T$ with $\mathcal{I}(\cdot)$ denote the imaginary part of the argument. The Fisher information matrix for $\tilde{\boldsymbol{\psi}}$ is defined as [7]

$$\mathbf{J}(\tilde{\boldsymbol{\psi}}) = \mathrm{E}_{\tilde{\boldsymbol{\psi}}} \left[\frac{\partial}{\partial \tilde{\boldsymbol{\psi}}} \tilde{\Lambda}(\boldsymbol{\psi}; Y(t)) \left(\frac{\partial}{\partial \tilde{\boldsymbol{\psi}}} \tilde{\Lambda}(\boldsymbol{\psi}; Y(t)) \right)^H \right], \tag{A.10}$$

with $(\cdot)^H$ denoting hermitian transposition. Using (A.3) and (A.5), we obtain after some algebraic manipulations

$$\mathbf{J}(\tilde{\boldsymbol{\psi}}) = \gamma \frac{8\pi^2}{E_\mathrm{p}} \mathcal{R}\left\{ \mathbf{G}^\mathbf{H} \mathbf{M}(\tilde{\boldsymbol{\psi}}) \mathbf{G} \right\} \tag{A.11}$$

where

$$\mathbf{G} = \mathrm{diag}\{\mathbf{N}, \mathbf{1}, \mathbf{1}, \mathbf{1}\}$$

$$\mathbf{M}(\tilde{\boldsymbol{\psi}}) = \int \begin{bmatrix} |\alpha|^2\Delta f^2 & -|\alpha|^2 t\Delta f & -\jmath\alpha\frac{\Delta f}{2\pi} & -\alpha\frac{\Delta f}{2\pi} \\ |\alpha|^2 t\Delta f & |\alpha|^2 t^2 & \jmath\alpha t & \alpha t \\ \jmath\alpha\frac{\Delta f}{2\pi} & -\jmath\alpha t & 1 & 0 \\ \alpha\frac{\Delta f}{2\pi} & -\alpha t & 0 & 1 \end{bmatrix} \otimes \mathbf{C}(\boldsymbol{\theta}, \boldsymbol{\theta}, t)\mathrm{d}t$$

with diag{} stands for a block diagonal matrix with the column vector on its diagonal, $\otimes$ denoting the Kronecker product, $\mathbf{1}$ being a column vector of all ones, $\mathbf{N} = [n(1), n(2), \ldots, n(NK)]^T$, and $\mathbf{C}(\boldsymbol{\theta}, \boldsymbol{\theta}^{'}, t)$ denoting the $NK \times NK$ matrix with (i, j)th entry $s_i(t, \boldsymbol{\theta})1(i \in \mathbf{I}_\mathrm{p})s_j^*(t, \boldsymbol{\theta}^{'})1(j \in \mathbf{I}_\mathrm{p})$.

The mth diagonal element of the inverted Fisher information matrix is the CRB on the variance of the estimation error of an unbiased estimator of $[\boldsymbol{\psi}_0]_m$. In particular,

$$\mathrm{CRB}_\tau = [\mathbf{J}^{-1}(\boldsymbol{\psi}_0)]_{1,1} \text{ and } \mathrm{CRB}_\nu = [\mathbf{J}^{-1}(\boldsymbol{\psi}_0)]_{2,2}. \tag{A.12}$$

The CRB is "local bound" in the sense that it depends essentially on the curvature of the main-lobe of the ambiguity function [7, Ch.10]. The narrower the main-lobe, the lower the CRB. As the SNR γ is a common factor that can be factored out from the Fisher information matrix (A.11), it is irrelevant when comparing the CRBs for various pilot signal selections. For such comparison, we can therefore consider the re-scaled versions $\gamma\mathrm{CRB}_\tau$ and $\gamma\mathrm{CRB}_\nu$.

3.3 Constrained Optimization Problem for Pilot Signal Design

It is well-known that the mean-squared error (MSE) of a nonlinear estimator such as (A.7) exhibits a so-called threshold effect [7]: If the SNR drops below a certain threshold value γ^{th}, there is an abrupt increase in the MSE of the estimator. We define γ^{th} for our particular application as follows:

Definition 1
The threshold value of a nonlinear estimator of (τ_0, ν_0) that asymptotically approaches the CRBs in (A.12) as the SNR increases is

$$\gamma^{th} = \max\{\gamma_{\hat{\tau}}^{th}, \gamma_{\hat{\nu}}^{th}\}$$

with

$$\gamma_{\hat{\tau}}^{th} = \min\{\gamma^{'} : \mathrm{MSE}_{\hat{\tau}}(\gamma) \leq 2\mathrm{CRB}_\tau(\gamma) \text{ for all } \gamma > \gamma^{'}\},$$

$$\gamma_{\hat{\nu}}^{th} = \min\{\gamma^{'} : \mathrm{MSE}_{\hat{\nu}}(\gamma) \leq 2\mathrm{CRB}_\nu(\gamma) \text{ for all } \gamma > \gamma^{'}\}.$$

The threshold effect is caused by outliers which occur if the estimate move from the main-lobe of the ambiguity function in (A.9) to one of its side-lobes

due to noise. The probability that outliers occur at a particular SNR is closely connected to the magnitude of the highest side-lobe of the normalized ambiguity function [7] [2]. We define the normalized side-lobe level (NSL) as the magnitude of the highest side-lobe of the normalized ambiguity function. A high NSL leads to a high sensitivity of the estimator towards noise, therefore, leading to high γ^{th}. Determining γ^{th} requires time-consuming Monte Carlo simulations. As an alternative, we can numerically obtain the NSL with a much lower computational effort.

To keep γ^{th} low, the NSL needs to be minimized. At the same time, to minimize the estimation error when the SNR is larger than γ^{th}, the CRBs also need to be minimized. But there is a tradeoff between the NSL and the CRBs. To account for this tradeoff, we formulate the design of the pilot signal as a constrained optimization problem[2]:

$$\begin{aligned} &\arg\min_{\mathbf{I}_\mathrm{p}\in\mathbf{I}} \mathrm{NSL}(\mathbf{I}_\mathrm{p}) \\ &\text{subject to } |\mathbf{I}_\mathrm{p}| = N_\mathrm{p} \\ &\mathrm{CRB}_\tau(\mathbf{I}_\mathrm{p}) < \mathrm{CRB}_\tau(\mathbf{I}_0), \\ &\mathrm{CRB}_\nu(\mathbf{I}_\mathrm{p}) < \mathrm{CRB}_\nu(\mathbf{I}_0), \end{aligned} \tag{A.13}$$

where $\mathbf{I}_0$ is a reference pilot signal with $|\mathbf{I}_0| = N_\mathrm{p}$.

The optimization procedure (A.13) differs from the optimization procedure formulated in [13] in three respects. First, whereas both procedures make use of the NSL as the first figure of merit, the former utilizes the CRBs as the second figure, while the latter utilizes the 3 dB bandwidth. Note that the CRBs can be easily computed via (A.12) and the NSL can be computed numerically. Second, while the latter procedure accounts for the tradeoff between the two figures of merit by specifying a weighted sum of them as the objective function to be optimized, the former deals with this tradeoff by means of a constrained optimization. Third, the procedure in [13] is constrained to the delay domain only, while (A.13) extends over the delay-Doppler domain.

[2] We could equally formulate another optimization problem which takes the NSL as constraint and minimizes the CRBs, i.e.

$$\begin{aligned} &\arg\min_{\mathbf{I}_\mathrm{p}\in\mathbf{I}} \mathrm{CRB}_\tau(\mathbf{I}_\mathrm{p}), \mathrm{CRB}_\nu(\mathbf{I}_\mathrm{p}) \\ &\text{subject to } |\mathbf{I}_\mathrm{p}| = N_\mathrm{p} \\ &\mathrm{NSL}(\mathbf{I}_\mathrm{p}) < \mathrm{NSL}(\mathbf{I}_0). \end{aligned}$$

In this case, however, one needs to simultaneously optimize two conflicting objectives (CRB_τ and CRB_ν). Indeed, reducing CRB_τ might increase CRB_ν and vice versa.

4 A Genetic Algorithm for Pilot Signal Design

The global optimal solution to the above optimization problem may be in principle found by exhaustive search. However, this search is unfeasibly complex since the number of possible patterns is $\binom{NK}{N_\mathrm{p}}$, which is large even for moderate values of NK and N_p. A feasible alternative is to use a genetic algorithm. Genetic algorithms are easy to implement, have fast convergence and are able to avoid local extrema [11]. Although the obtained solutions are suboptimal, genetic algorithms are well-suited for combinatorial optimization problems. We refer the interested reader to [11] for the basics and the applications of such algorithms in signal processing.

We propose the genetic algorithm described below (Algorithm 1) to solve the constraint optimization problem (A.13). In this context, we define the "chromosomes" to be the pilot patterns. The algorithm can be conveniently extended to jointly design the pilot pattern and the pilot signature by additionally including the complex amplitudes in each chromosome.

5 Numerical Performance Evaluation

In this section, we utilize the proposed Algorithm 1 to design pilot signals for delay-Doppler estimation and then compare their performance to state-of-the-art pilot signals via Monte Carlo simulations of the MSE for the joint delay-Doppler shift estimator in (A.7). For these investigations, we use the settings summarized in Table A.1. The considered OFDM frame corresponds to 24 resource blocks according to the LTE specifications.

In Fig. A.2, we consider four pilot signals (a)-(d): Pattern (a) is the equispaced and equipowered pilot signal; Pattern (b) is the "perfect periodic" Costas array; Pattern (c) is the pilot pattern designed using Algorithm 1; and Pattern (d) is obtained using Algorithm 1 modified to design pilot pattern and signature jointly. Fig. A.2 reports patterns (a)-(d), along with their associated magnitude of the ambiguity functions, NSLs, and CRBs. From the results in Fig. A.2, we make two observations: Firstly, we observe that the pilot signal designed with Algorithm 1, i.e. (c) and (d), leads to much lower fitness parameters (NSL and CRBs) than that are obtained for (a) and (b). Thus, as expected, Algorithm 1 is able to improve the design of pilot signals in terms of their fitness parameters as expected. We remark that the noticeably high NSL of the ambiguity function associated with the "perfect periodic" Costas array (b) is induced jointly by the high side-lobes near the main-lobe of the ambiguity function of the Costas array and the repetition of this array with the selected spacing in the frequency domain. Secondly, we observe only small differences between the NSLs and the CRBs for pilot signals (c) and (d). This observation indicates that the impact of the pilot pattern is predominant on

Algorithm 1: Genetic algorithm for the design of pilot pattern for joint delay-Doppler estimation.

←: assignment operation. URWR: uniformly at random without replacement.

Initialization: Set N_{ind}, N_{elite}(even number), N_{m} and randomly generate the initial population $\mathbf{I}(0) = \{\mathbf{I}_{p_1}, \ldots, \mathbf{I}_{p_{N_{\text{ind}}}}\}$ with $|\mathbf{I}_{\text{p}_\text{i}}| = N_p$;

for $g = 0, 1, \ldots, MaxGen$ **do**

Elite selection: Form $\mathbf{I}_{\text{elite}}(g) \subset \mathbf{I}(g)$ consisting of the N_{elite} pilot patterns with the lowest fitness

$$F(\mathbf{I}_{\text{p}_\text{i}}) = \begin{cases} \text{NSL} & \text{CRB}_\tau(\mathbf{I}_{\text{p}_\text{i}}) < \text{CRB}_\tau(\mathbf{I}_0) \ \ \& \\ & \text{CRB}_\nu(\mathbf{I}_{\text{p}_\text{i}}) < \text{CRB}_\nu(\mathbf{I}_0) \quad ; \\ 1 & \text{otherwise.} \end{cases}$$

$\mathbf{I}(g+1) \leftarrow \mathbf{I}_{\text{elite}}(g)$;

for $j = 1, \ldots, \frac{N_{\text{elite}}}{2}$ **do**

Pick two elements $\mathbf{I}^{'}, \mathbf{I}^{''} \in \mathbf{I}_{\text{elite}}(g)$ URWR ;

Generate offspring $\mathbf{I}_{\text{off}} \subset \mathbf{I}^{'} \cup \mathbf{I}^{''}$ by picking N_{p} elements from $\mathbf{I}^{'} \cup \mathbf{I}^{''}$ URWR ;

Mutation: Pick N_{m} elements from $\mathbf{I}_{\text{off}}$ URWR and substitute them by N_{m} elements picked from $\mathbf{I}^{c}_{\text{off}} = \mathbf{I} \setminus \mathbf{I}_{\text{off}}$ URWR to generate $\mathbf{I}_{\text{m}}$;

$\mathbf{I}_{\text{elite}}(g) \leftarrow \mathbf{I}_{\text{elite}}(g) \setminus \{\mathbf{I}^{'}, \mathbf{I}^{''}\}$;

Update population: $\mathbf{I}(g+1) \leftarrow \mathbf{I}(g+1) \cup \mathbf{I}_{\text{m}}$;

end

end

Table A.1: Simulation Settings

OFDM system:
$N = 288$, $K = 7$, $N_{\mathrm{p}} = 96$
$T_{\mathrm{p}} = T + T_{\mathrm{cp}} = 66.7 + 6.67 = 73.4\,\mu\mathrm{s}$
Genetic algorithm:
$N_{\mathrm{ind}} = 100$, $N_{\mathrm{elite}} = 40$, $N_{\mathrm{m}} = 1$, $MaxGen = 80$
$\mathbf{I}_0$ = "Perfect periodic" Costas array
Estimation range $\tau \in [-\frac{T_{\mathrm{p}}}{2}, \frac{T_{\mathrm{p}}}{2}]$, $\nu \in [-\frac{\Delta f}{2}, \frac{\Delta f}{2}]$
Pilots are equipowered with zero phase unless otherwise specified.

the estimator performance compared to the impact of the pilot signature.

We now evaluate the reduction of MSE that can be obtained for the pilot signals designed with Algorithm 1 by means of Monte Carlo simulations using the joint delay-Doppler shift estimator (A.7). We demonstrate that although Algorithm 1 is proposed under simplified conditions, the designed pilot signals are also appropriate under more realistic conditions. We consider three scenarios of increasing realism: Scenario 1 is the single-path propagation with pilot-only transmission, i.e. the scenario for which estimator (A.7) coincides with the maximum likelihood estimator of the delay and Doppler shift. Scenario 2 is the same as Scenario 1, but includes data transmission. Scenario 3 is with both data transmission and multipath propagation. In the first two scenarios, we assume without loss of generality that $\psi_0 = [0, 0, 1]^T$.

5.1 Scenario 1: Single-Path Propagation, Without Data Transmission

Fig. A.3 reports the MSE of estimator (A.7) computed from Monte Carlo simulations, using pilot signals (a)-(c) in Fig. A.2. It appears that pilot signal (c) designed with Algorithm 1 leads to a threshold gain of 7 dB and 2 dB compared to pilot signals (a) and (b) respectively, as a result of the significant reduction of the NSL.

5.2 Scenario 2: Single-Path Propagation, With Data Transmission

So far the effect of data signals on the estimation performance of the estimator (A.7) has been neglected in the literature, see e.g. [10] [5] [6] [13]. In this subsection, we compare the effect of data signals on patterns (b) and (c).

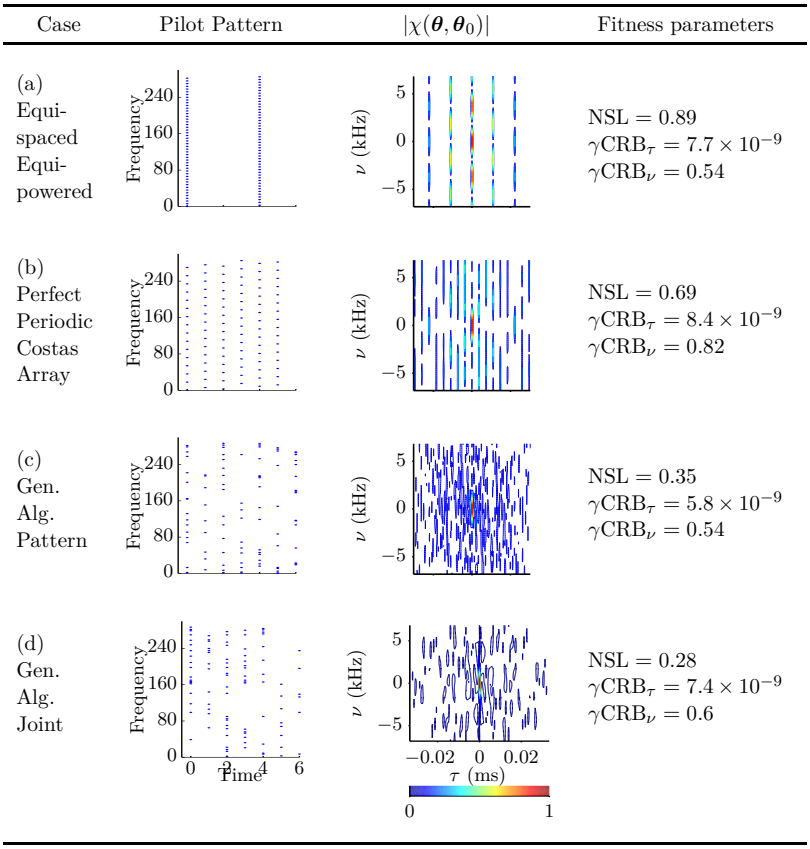


Fig. A.2: The considered pilot signals. The pilot pattern and signature of the pilot signal in panel (d) are jointly optimized: the amplitudes and phases of all pilots are drawn independently according to a uniform distribution on $[0, 1]$ and $[0, 2\pi]$ respectively during the initialization of Algorithm 1. At each mutation stage of the algorithm, the signature of the N_m selected pilots is drawn similarly. After each random drawing, the pilot signature is scaled such that its energy equals 1.

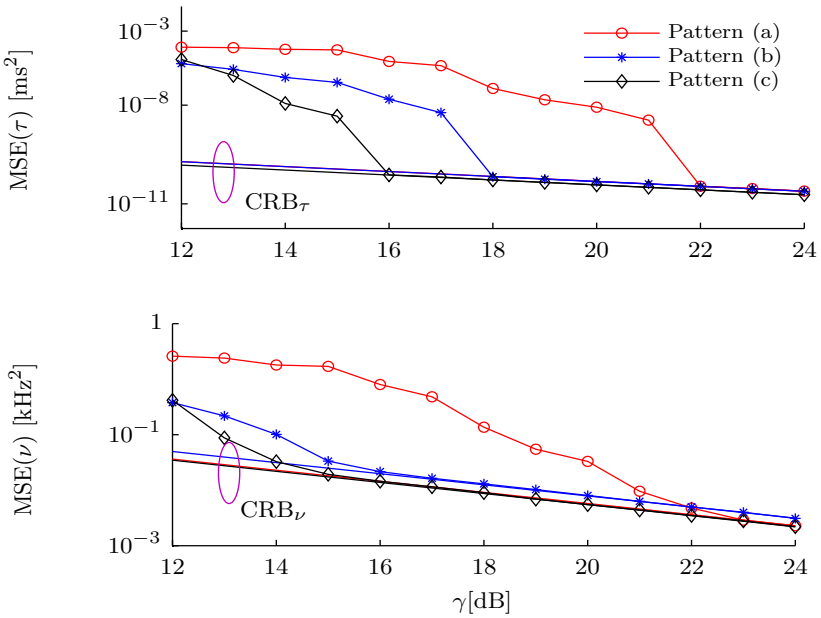


Fig. A.3: Scenario 1: MSE performance of estimator (A.7) and CRBs versus SNR when using pilot patterns (a), (b), and (c) in Fig. A.2. The corresponding threshold values (γ^{th}) are 23 dB, 18 dB and 16 dB, respectively. Each point is obtained from 10000 Monte Carlo trials.

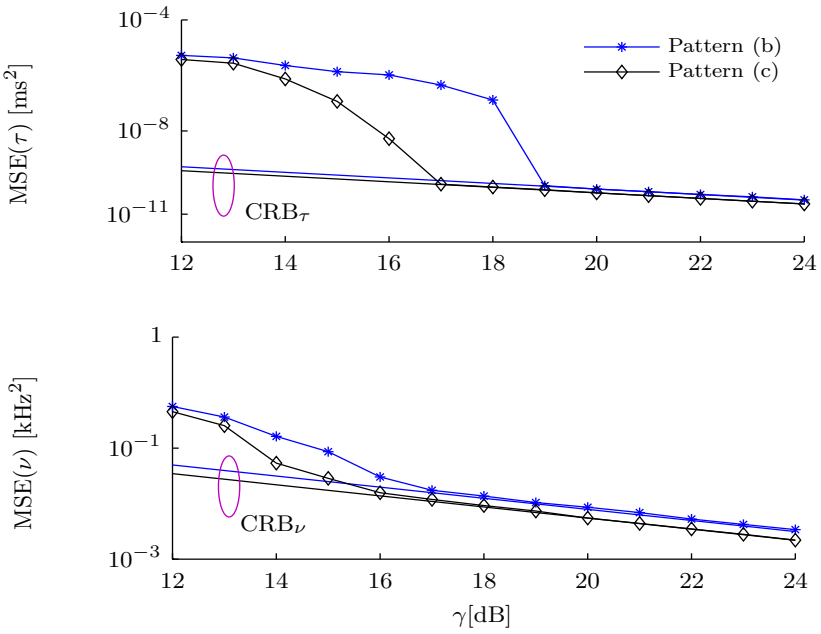


Fig. A.4: Scenario 2: MSE performance of estimator (A.7) and CRBs versus SNR when using pilot patterns (b) and (c) in Fig. A.2. Each point is obtained from 10000 Monte Carlo trials.

During the data transmission phase, (A.3) reads $Y(t) = \alpha_0(s_\mathrm{p}(t;\boldsymbol{\theta}_0) + s_\mathrm{d}(t;\boldsymbol{\theta}_0)) + N(t)$. The objective function in (A.7) is given by

$$\begin{aligned} Z(\boldsymbol{\theta};Y(t)) &= \left|\int s_\mathrm{p}^*(t;\boldsymbol{\theta})Y(t)\mathrm{d}t\right|^2 \\ &= \Big|\alpha_0\chi(\boldsymbol{\theta},\boldsymbol{\theta}_0) + \underbrace{\alpha_0\int s_\mathrm{p}^*(t;\boldsymbol{\theta})s_\mathrm{d}(t;\boldsymbol{\theta}_0)\mathrm{d}t}_{\text{Interference}} \\ &\qquad + W(\boldsymbol{\theta})\Big|^2. \end{aligned} \tag{A.14}$$

Fig. A.4 reports the MSE of estimator (A.7) when using pilot signals (b) and (c). A comparison with the MSE results reported Fig. A.3 shows that the interference caused by data transmission only affects the threshold value, which is shifted to the right by approximately 1 dB. In the high SNR regime, the data interference has no significant effect on the estimator performance.

5.3 Scenario 3: Multipath Propagation, With Data Transmission

In the third and most realistic scenario, we consider both data transmission and multipath propagation. In this scenario, the objective function (A.7) reads

$$Z(\boldsymbol{\theta}; Y(t)) = \Big| \sum_{l=0}^{L-1} \alpha_l \chi(\boldsymbol{\theta}, \boldsymbol{\theta}_l) + \underbrace{\sum_{l=0}^{L-1} \alpha_l \int s_{\mathrm{p}}^*(t; \boldsymbol{\theta}) s_{\mathrm{d}}(t; \boldsymbol{\theta}_l) \mathrm{d}t}_{\text{Interference}} + W(\boldsymbol{\theta}) \Big|^2. \tag{A.15}$$

First we consider the case where the main-lobe and the side-lobes of the ambiguity function of the pilot signal devised with Algorithm 1 are respectively sufficiently narrow and low enough, so that the L multipath components can be resolved in the delay-Doppler domain. This implies that, if one discards the effect of noise and data interference in (A.15), $Z(\boldsymbol{\theta}; Y(t))$ exhibits L well-separated dominant peaks, each peak being contributed by one multipath component. Each peak uniquely corresponds to the main-lobe of one of the weighted ambiguity function in the first summand in (A.15). In this case, the joint ML estimator of the L pairs $\{(\tau_l, \nu_l)\}_{l=1}^{L}$ is accurately approximated by L independent ML estimators, one for each pair. The outputs of these L estimators are the L delay-Doppler arguments corresponding to the L largest maxima of $Z(\boldsymbol{\theta}; Y(t))$. The same approximation holds true in the presence of data interference and of noise at high SNR and even at medium SNR since the NSL is low. Thus, in a scenario with well-separable multipath components, a pilot signal designed using Algorithm 1 still essentially keeps its optimality properties in medium and high SNR regime.

A necessary condition for multipath components to be separable is that the bandwidth of the OFDM system is large enough. Due to practical constraints on the available bandwidth of the OFDM system and the number of pilots, not all path components may be resolved in $Z(\boldsymbol{\theta}; y(t))$ in (A.15). We consider such a case in the following and show that the pilot pattern designed using Algorithm 1 is still a good choice. We use (A.7) to obtain the delay and Doppler shift estimates and compute the MSE of the delay and Doppler shift by using the first path component as the reference. As indicated in the introduction, this estimate can be used—in combination with other such estimates computed from other transmission links—for localization in a TOA or TDOA based positioning method, or also for synchronization [8]. However, due to the unresolvable path components with higher delay than the first component, the estimator is expected to be biased.

The "Extended Vehicular A" channel model specified in the 3GPP LTE

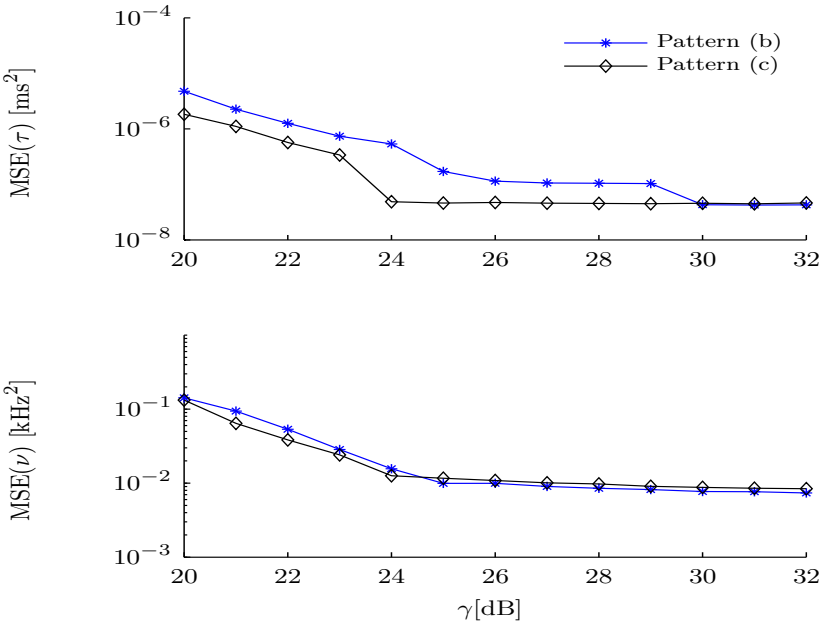


Fig. A.5: Scenario 3: MSE performance of estimator (A.7) versus SNR when using pilot patterns (b) and (c) in Fig. A.2. Each point is obtained from 10000 Monte Carlo trials.

standard [1] is used to generate a new channel impulse response for each simulation trial. The values of the delay, Doppler shift and weight magnitude of the $L = 9$ multipath components are kept fixed while generating the channel responses. The phases of the path weights are drawn independently from a uniform distribution on $[0, 2\pi)$ for each trial.

A comparison of the MSE curves depicted in Fig. A.5 with those reported in Fig. A.4 shows that a much higher error floor appears at high SNR due to bias caused by unresolved multipath components. In addition, the multipath channel leads to a significantly shifts of the thresholds. Specifically, for the pilot signal designed with Algorithm 1, a pronounced threshold appears at 25 dB while for the "perfect periodic" Costas array, it appears at 30 dB. This observation indicates that even though pilot pattern (c) is optimized for the idealized scenario ignoring multipath propagation and data transmission, it still leads to better performance than obtained by using the "perfect periodic" Costas array.

6 Conclusion

The conventional equispaced and equipowered pilot signals, as used in LTE, is suboptimal for joint delay and Doppler estimation. It has an ambiguity function with a high normalized side-lobe level (NSL), which causes the correlator-based estimator (A.7) to exhibit a high threshold value. The proposed genetic algorithm generates pilot signals that minimize the NSL, while maintaining the CRBs for the delay and Doppler shift estimation below a target value. Compared to the "perfect periodic" Costas arrays, these pilot signals produce much lower NSL and CRBs. The results clearly exemplify that the possible reduction in MSE can be achieved with the same number of pilots. An additional finding is that the pilot pattern affects more significantly the NSL and the CRBs than the pilot signature does. Our genetic algorithm can generate close-to-optimal pilot signals regardless of the OFDM frame size and the number of pilots. This computation can be done offline. We also show that the pilot signals computed with the genetic algorithm remain a good choice in single-path and multipath propagation conditions during the data transmission phase when a correlator is employed for delay and Doppler shift estimation.

Among the interesting open research avenues, we would like to mention the extension of the constrained optimization problem to account for transmission across multipath channels, especially when more sophisticated channel estimators are used. The performance of these estimators could then be assessed by means of performance criteria traditionally employed in communications, such as channel estimation error and bit-error-rate.

Acknowledgment

This work has been funded by the project ICT-248894 Wireless Hybrid Enhanced Mobile Radio Estimators–Phase 2 (WHERE2).

References

[1] 3GPP. Base station (BS) radio transmission and reception. 3GPP TS 36.104, V8.12.0 (2011-06) Technical Specification, 2011.

[2] F. Athley. Threshold region performance of maximum likelihood direction of arrival estimators. *IEEE Trans. Signal Process.*, 53(4):1359–1373, Apr. 2005.

[3] P. Chestnut. Emitter location accuracy using TDOA and differential doppler. *IEEE Trans. Aerosp. Electron. Syst.*, AES-18(2):214–218, Mar. 1982.

[4] J. Costas. A study of a class of detection waveforms having nearly ideal range-Doppler ambiguity properties. *Proc. IEEE*, 72(8):996–1009, Aug. 1984.

[5] A. Dammann, C. Mensing, and S. Sand. On the benefit of location and channel state information for synchronization in 3GPP-LTE. In *European Wireless Conf.*, pages 711–717, 2010.

[6] J.-C. Guey. Synchronization signal design for OFDM based on time-frequency hopping patterns. In *IEEE International Conf. on Commun.*, pages 4329–4334, June 2007.

[7] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory - Part III*. John Wiley & Sons, 2001.

[8] C. Mensing. *Location Determination in OFDM Based Mobile Radio Systems*. PhD thesis, Technische Universität München, 2013.

[9] H. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1994.

[10] A. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location. *IEEE Signal Process. Mag.*, 22(4):24–40, 2005.

[11] K. Tang, K. Man, S. Kwong, and Q. He. Genetic algorithms and their applications. *IEEE Signal Process. Mag.*, 13(6), Nov. 1996.

[12] L. Tong, B. Sadler, and M. Dong. Pilot-assisted wireless transmissions: general model, design criteria, and signal processing. *IEEE Signal Process. Mag.*, 21(6):12–25, Nov. 2004.

[13] O. Ureten, S. Tascioglu, N. Serinken, and M. Yilmaz. Search for OFDM synchronization waveforms with good aperiodic autocorrelations. In *Canadian Conf. on Elect. and Compt. Eng.*, volume 1, pages 13–18, May 2004.

[14] P. Woodward. *Probability and Information Theory, with Applications to Radar*. Pergamon Press, 1953.

Paper B

Direct Ranging in Multi-path Channels Using OFDM Pilot Signals

Lishuai Jing, Troels Pedersen, Bernard H. Fleury

The paper has been published in the
15th IEEE International Symposium on Signal Processing Advances in Wireless Communications (SPAWC)
June 22 - June 25, Toronto, Canada, 2014.

The layout has been revised.

Abstract

OFDM ranging is becoming important for positioning using terrestrial wireless networks. Conventional ranging methods rely on a two-step approach: range related parameters, such as the time of arrival (TOA), the bias induced by non-line-of-sight (NLOS) propagations etc., are first estimated, based on which the range is then inferred. In multi-path conditions, two-step range estimators which employ the correlator-based estimator or the energy detector lead to poor ranging accuracy when applied in non-ultra-wideband scenarios due to a bias. More advanced ranging schemes that estimate all multi-path components using a multidimensional search procedure provide higher ranging accuracy but have a prohibitive complexity. In this work, we propose a novel direct ranging technique that uses a point process formulated channel model. Based on this model, we derive an approximate maximum likelihood estimator of the range. In contrast to the estimator which requires a multidimensional search procedure, the proposed estimator does not demand the knowledge of the exact number of multi-path components and these components are separable. If the power delay spectrum of the multi-path channel and the signal-to-noise-ratio (SNR) are known, the complexity of the proposed estimator is tractable. We show by means of Monte Carlo simulations that this estimator outperforms the correlator-based estimator.

1 Introduction

Accurate localization is becoming important for terrestrial wireless systems, in particular for OFDM systems such as WLAN, LTE and its extension LTE-A [14] [11]. One approach to improve the localization accuracy is to design high precision ranging techniques [1] [15]. State-of-the-art ranging techniques follow a two-step approach. First, parameters, such as the received signal strength, the TOA, the bias induced by NLOS propagations etc., are estimated from the received signal. Then, these estimates are used for ranging [1]. Since some of these information bearing parameters are readily available in communication systems, two-step ranging methods are very popular.

Two-step approaches employing OFDM signals have been considered in [15] [2] [16]. Wang *et al.* [16] proposed a maximum-likelihood ranging method based on OFDM signals for a scenario with separable multi-path components. Due to the assumed separability (in the delay domain) of these components, the obtained estimator converges to the correlator-based estimator [1]. However, for the separability condition to hold, a large system bandwidth is needed and even in this case it is not guaranteed that all paths are separable. In addition, because this method relies on the detection and estimation of the line-of-sight (LOS) path, it is sensitive to fading of early non-separable com-

ponents [1]. MUltiple SIgnal Classification (MUSIC) algorithm was applied by Zhao *et al.* in [2] to estimate the delays of all multi-path components assuming their exact number is known. In [15], Wang *et al.* derived the Cramér-Rao bound (CRB) on the mean square error (MSE) of the range estimator and investigated how the OFDM signal parameters and the spacing between the multi-path components affect the bound. The CRB is a lower bound for the MSE of the estimator proposed in [13], which requires estimation of the delays of all separable multi-path components. The bound in [15] and the methods in [2] [16] [13] require the separability of the multi-path components and the knowledge of their exact number, which is generally difficult to estimate reliably. Furthermore, the required multidimensional search is impractical for a realistic number of multi-path components. As an alternative to the two-step approach, ranging can be performed in one step—referred to as "direct" ranging—avoids the need for the detection of the first-path and the estimation of its parameters. Despite the ability to bypass both the first-path detection problem and the path separability requirement, direct ranging has attracted little attention in the literature.

In this contribution, we address the problem of direct ranging using OFDM pilot signals in multi-path channels. The objective is to obtain a ranging estimator with low complexity, which does not rely on first-path detection, any separability condition, and the knowledge of the number of multi-path components. To that end, we formulate a multi-path channel model using a point process approach [3] [5] [9]. The channel transfer function is reformulated such that the range parameter is factored out to make it accessible for direct estimation. We then propose a direct ranging method using a Gaussian approximation of the channel transfer function. The method avoids the requirement of knowing the exact number of multi-path components and relaxes the constraint on their separability. Given the SNR and the RMS delay spread of the channel, the proposed estimator is computationally tractable. Simulation results demonstrate that the proposed estimator outperforms the correlator-based estimator.

2 System Model

We consider a single-input single-output OFDM setup with N active sub-carriers. An OFDM symbol with time duration T is generated by multiplexing a sequence of data symbols and known pilot symbols onto N orthogonal sub-carriers. The adjacent sub-carrier spacing is defined as $\Delta f = \frac{1}{T}$. A cyclic prefix with duration T_{cp} is appended to prevent inter-symbol and inter-carrier interference. We index the N active sub-carriers with the set $\mathbf{I} = \{1, 2, \ldots, N\}$. Of these sub-carriers, $N_{\text{p}} = |\mathbf{I}_{\text{p}}|$ are pilots indexed by $\mathbf{I}_{\text{p}} \subseteq \mathbf{I}$.

We address estimation of the range parameter d based on the pilot signals. The multi-path channel is assumed to be time-invariant during the transmission

of each OFDM symbol. Removing the cyclic prefix and concatenating the received pilot signals in the observation vector $\mathbf{y}$, we obtain the signal model in the frequency domain:

$$\mathbf{y} \quad = \quad \mathbf{A}\mathbf{h}(d) + \mathbf{w}, \tag{B.1}$$

where $\mathbf{A} = \text{diag}\{a_n : n \in \mathbf{I}_\text{p}\}$ is a diagonal matrix with a_n denoting the nth pilot symbol, the vector $\mathbf{h}(d) = [h(d; n\Delta f) : n \in \mathbf{I}_\text{p}]$, contains the samples of the channel frequency response, and $\mathbf{w}$ is a white circular-symmetric complex Gaussian noise vector with component variance σ^2. We define the SNR as $\frac{E_s}{\sigma^2}$ with $E_s = E[|a_n|^2]$.

The channel frequency response is modeled as a sum of delayed and attenuated multi-path components [9]:

$$h(d; f) = \underbrace{q\alpha_0 e^{-j2\pi f(\tau_0 + \frac{d}{c})}}_{\text{LOS term}} + \underbrace{\sum_{l=1}^{L} \alpha_l e^{-j2\pi f(\tau_l + \frac{d}{c})}}_{\text{Tail}}, \tag{B.2}$$

where path l has complex gain α_l and excess delay τ_l and c is the speed of light. The indicator q specifies the settings of the LOS path component. For LOS channels i.e. $q \neq 0$, q adjusts the power of the LOS component. When $q = 0$, the system operates in NLOS conditions. The delay of the LOS path is $\frac{d}{c}$ and thus we set the LOS excess delay equal to zero: $\tau_0 = 0$. The excess delays of the NLOS paths form a point process $\mathcal{T} = \{\tau_1, \tau_2, \ldots\}$ with intensity function $\rho(\tau)$. Note that the number of multi-path components L, i.e. the cardinality of $\mathcal{T}$, is not necessarily deterministic or finite under such channel formulation. We also assume that

$$E[\alpha_l|\tau_l] = 0, \;\; E[\alpha_l \alpha_{l'}^*|\tau_l, \tau_{l'}] = \begin{cases} \sigma_\alpha^2(\tau_l), & l = l' \\ 0, & \text{otherwise.} \end{cases} \tag{B.3}$$

For convenience, we reformulate (B.2) as the product of a range-dependent factor $r(d; f)$ and a factor $\varepsilon(f)$ independent of d:

$$h(d; f) = \underbrace{e^{-j2\pi f \frac{d}{c}}}_{r(d;f)} \underbrace{\left(q\alpha_0 + \sum_{\tau_l \in \mathcal{T}} \alpha_l e^{-j2\pi f \tau_l} \right)}_{\varepsilon(f)}. \tag{B.4}$$

The assumption that $\varepsilon(f)$ is independent of d is a simplification which may or may not be realistic. Here, we employ it to simplify the forthcoming derivations. We leave investigation of more sophisticated distance dependent channel models such as presented in [4] to future works. Defining the diagonal matrix $\mathbf{R}(d) =$

$\text{diag}\{r(d; n\Delta f) : n \in \mathbf{I}_\text{p}\}$ and the vector $\boldsymbol{\varepsilon} = [\varepsilon(n\Delta f) : n \in \mathbf{I}_\text{p}]^T$, the channel vector reads

$$\mathbf{h}(d) = \mathbf{R}(d)\boldsymbol{\varepsilon}. \tag{B.5}$$

Following the assumptions (B.3), $E[\boldsymbol{\varepsilon}] = \mathbf{0}$ and thus $E(\mathbf{y}) = \mathbf{0}$. With these results, the covariance matrix of the observation vector $\mathbf{y}$ is given by

$$\mathbf{C}_\mathbf{y}(d) = \mathbf{A}\mathbf{R}(d)\mathbf{C}_{\boldsymbol{\varepsilon}}\mathbf{R}^H(d)\mathbf{A}^H + \sigma^2\mathbb{I}, \tag{B.6}$$

where $\mathbf{C}_{\boldsymbol{\varepsilon}} = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^H]$ with $(\cdot)^H$ denoting conjugate transpose and $\mathbb{I}$ being the identity matrix. Inspired by [5], $\mathbf{C}_{\boldsymbol{\varepsilon}}$ can be computed from an underlying channel model. Entry (m, n) of $\mathbf{C}_{\boldsymbol{\varepsilon}}$ reads

$$[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} = q^2\sigma_\alpha^2(0) + E\left[\sum_{\tau_l,\tau_{l'}\in\mathcal{T}} \alpha_l\alpha_{l'}^* e^{-j2\pi\Delta f(m\tau_l - n\tau_{l'})}\right].$$

By the law of total expectation, conditioning on the point process $\mathcal{T}$, and utilizing (B.3), we obtain

$$[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} = q^2\sigma_\alpha^2(0) + E\left[\sum_{\tau_l\in\mathcal{T}} \sigma_\alpha^2(\tau_l) e^{-j2\pi\Delta f(m-n)\tau_l}\right].$$

Applying Campbell's theorem [7] yields

$$\begin{aligned}[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} &= \int_0^\infty \underbrace{\sigma_\alpha^2(\tau)(\rho(\tau) + q^2\delta(\tau))}_{P(\tau)} e^{-j2\pi(m-n)\Delta f\tau}\mathrm{d}\tau \\ &= \mathcal{F}\{P(\tau)\}((m-n)\Delta f),\end{aligned} \tag{B.7}$$

where $\delta(\cdot)$ denotes the Dirac delta function, $P(\tau)$ is the delay power spectrum (the average delay power profile) and $\mathcal{F}$ denoting the Fourier transform [5]. In practice, the delay power spectrum can be evaluated empirically or approximated using an appropriated channel model [12] [10]. We further assume that $\rho(\tau) = \rho$ with ρ being a constant. Thus, we assume a constant arrival rate for the delays induced in the "Tail". With this assumption, (B.7) reads

$$[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} = q^2\sigma_\alpha^2(0) + \rho\underbrace{\int_0^\infty \sigma_\alpha^2(\tau) e^{-j2\pi(m-n)\Delta f\tau}\mathrm{d}\tau}_{g_{mn}=\mathcal{F}\{\sigma_\alpha^2(\tau)\}((m-n)\Delta f)}. \tag{B.8}$$

To gain some insight into the impact of the properties of the channel model on $\mathbf{C}_{\boldsymbol{\varepsilon}}$, we consider the following three example models.

Example 2.1
The number L is fixed and $\sigma_\alpha^2(\tau) = \frac{1}{L}$. The delays in the "Tail" are drawn independently and uniformly on $[0, T_{\text{cp}}]$. In this case, $\mathcal{T}$ is a Binomial point process. Hence, $\rho = \frac{L}{T_{\text{cp}}}$. Consequently, (B.8) reads

$$[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} = q^2 \frac{1}{L} + \text{sinc}((m-n)\Delta f T_{\text{cp}}) e^{-j\pi(m-n)\Delta f T_{\text{cp}}} \tag{B.9}$$

with $\text{sinc}(x) = \sin(\pi x)/(\pi x)$. In the LOS scenario, the covariance matrix depends on the exact number of paths L, which is generally unknown in practice. In the NLOS scenario, i.e. $q = 0$, the covariance matrix looses the dependency on L since the first term in (B.9) vanishes due to the somewhat artificial assumption $\sigma_\alpha^2(\tau) = \frac{1}{L}$. Note that the involved assumptions are similar to those used to derive the robust Wiener filter [8].

Example 2.2
The number L is fixed and motivated by experimental observations [4], we assume that

$$\sigma_\alpha^2(\tau) = C \exp(-\frac{\tau}{\lambda}), \tag{B.10}$$

where C is a positive constant and λ denotes the RMS delay spread of the "Tail" of the multi-path channel. We reuse the assumptions invoked in Example 2.1 except the assumption on $\sigma_\alpha^2(\tau)$. Assuming that $\int_{T_{\text{cp}}}^{\infty} \sigma_\alpha^2(\tau) \text{d}\tau$ is negligible, (B.8) reads

$$[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} \quad = \quad q^2 C + \frac{L}{T_{\text{cp}}} g_{mn}, \tag{B.11}$$

where

$$g_{mn} = C \frac{1 - e^{-(j2\pi(m-n)\Delta f + \frac{1}{\lambda})T_{\text{cp}}}}{j2\pi(m-n)\Delta f + \frac{1}{\lambda}}. \tag{B.12}$$

Notice that the covariance matrix depends on L.

Example 2.3
$\mathcal{T}$ is modeled as a homogeneous Poisson point process on $[0, T_{\text{cp}}]$ with rate ρ and exponential power decay for $\sigma_\alpha^2(\tau)$. This is a special case of Turin's model [3]. Then L is a Poisson random variable with mean $\mu_L = E[L] = \rho T_{\text{cp}}$. Assuming that $\int_{T_{\text{cp}}}^{\infty} \sigma_\alpha^2(\tau) \text{d}\tau$ is negligible and utilizing (B.10), (B.8) reads

$$[\mathbf{C}_{\boldsymbol{\varepsilon}}]_{mn} = q^2 C + \rho g_{mn} \tag{B.13}$$

with g_{mn} defined as in (B.12). We observe that the covariance matrix $\mathbf{C}_{\boldsymbol{\varepsilon}}$ does not depend on the exact number of paths of a specific channel realization but depends on the intensity ρ and λ. The intensity ρ and λ may be provided by an appropriate channel model.

3 Direct Maximum Likelihood Ranging Via Gaussian Approximations

The direct maximum likelihood estimator of d based on the observation $\mathbf{y}$ reads

$$\hat{d}_{ML} = \arg\max_{d} p(\mathbf{y}|d), \tag{B.14}$$

where $p(\mathbf{y}|d)$ denotes the likelihood function of d given $\mathbf{y}$. Estimator (B.14) is a "direct" range estimator since no intermediate parameters such as delays, complex gains, etc. are estimated. To implement (B.14), the likelihood function $p(\mathbf{y}|d)$ needs to be computed. In the considered case, however, $p(\mathbf{y}|d)$ is unknown. Instead, estimator (B.14) may be approximated as in [15] and [13] via a two-step approach. These methods, however, require the knowledge of the number of path components, which is generally unknown and hard to estimate.

Here, we follow the alternative approach of approximating $p(\mathbf{y}|d)$ with a Gaussian pdf $\tilde{p}(\mathbf{y}|d)$ with the same first- and second-order moments. This approximation is exact if $\boldsymbol{\varepsilon}$ is a Gaussian random vector. It is a reasonable approximation in a multi-path channel where L is large and $\sigma_\alpha^2(\tau)$ is a constant. In more realistic channels with an exponential power decay, the Gaussian approximation can be inaccurate. Since the first- and second-order moments of $\mathbf{y}$ are known by (B.6), this approximation leads to an estimator that can be derived analytically. Using $\tilde{p}(\mathbf{y}|d)$ instead of $p(\mathbf{y}|d)$ in (B.14) yields

$$\hat{d}_{AML} \quad = \quad \arg\max_{d} \ln \tilde{p}(\mathbf{y}|d), \tag{B.15}$$

where the log-likelihood $\ln \tilde{p}(\mathbf{y}|d)$ is of the form [6]

$$\ln \tilde{p}(\mathbf{y}|d) \propto -\ln\det(\mathbf{C}_{\mathbf{y}}(d)) - \mathbf{y}^H \mathbf{C}_{\mathbf{y}}^{-1}(d)\mathbf{y} \tag{B.16}$$

with $x \propto z$ denoting $x = z + constant$ and $\det(\cdot)$ denoting the determinant. Using the eigenvalue decomposition $\mathbf{C}_{\boldsymbol{\varepsilon}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, we can recast (B.6) as

$$\mathbf{C}_{\mathbf{y}}(d) = \mathbf{R}(d)\mathbf{G}\mathbf{R}^H(d),$$

with $\mathbf{G} = \mathbf{A}\mathbf{U}(\boldsymbol{\Lambda} + \mathbf{I}\sigma^2/E_s)\mathbf{U}^H\mathbf{A}^H$. Since $\mathbf{R}(d)$ is unitary, the determinant $\det(\mathbf{C}_{\mathbf{y}}(d)) = \det(\mathbf{G})$ does not depend on d and can be dropped. Thus,

$$\ln \tilde{p}(\mathbf{y}|d) \quad \propto \quad -\mathbf{y}^H \mathbf{R}(d)\mathbf{G}^{-1}\mathbf{R}^H(d)\mathbf{y}, \tag{B.17}$$

Table B.1: Simulation Settings

OFDM system: $N = 512$, $N_\mathrm{p} = 103$, $\Delta f = 15\,\mathrm{kHz}$, $T_\mathrm{p} = T + T_\mathrm{cp} = 66.7 + 5.4 = 72.1\,\mu\mathrm{s}$
Estimation range $d_\mathrm{obs} \in [0, 7\,\mathrm{km}]$; True range: $d = 1\,\mathrm{km}$. Results obtained from 10000 Monte Carlo trials are displayed.

where

$$\mathbf{G}^{-1} = \mathbf{A}\mathbf{U}(\mathbf{\Lambda} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}\mathbf{U}^H\mathbf{A}^H. \tag{B.18}$$

Since the matrices $\mathbf{AU}$ and $\mathbf{\Lambda}$ can be pre-computed and stored, the inversion of $\mathbf{G}$ amounts to compute the diagonal matrix $(\mathbf{\Lambda} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$. This circumvents the brute force inversion of $\mathbf{C}_\mathbf{y}(d)$ in (B.16) and thereby reduces the complexity of the estimator. We remark that $\mathbf{C}_{\boldsymbol{\varepsilon}}$ and thus $\mathbf{U}$ and $\mathbf{\Lambda}$ depend on the parameters of $P(\tau)$. It can be shown that the non-coherent correlator-based estimator [11] [1] [16] is a limiting case of the proposed estimator (B.15) when $q \to \infty$, which implies a single path channel.

4 Numerical Performance Evaluation

We first evaluate the performance of estimator (B.15) in a multi-path channel with different parameter settings and contrast it with the performance of the non-coherent correlator-based estimator [11] [16]. We omit the comparison with the energy detector, which is sensitive to the selected threshold value and provides inaccurate TOA estimates [1] [12]. We also omit the comparison with multidimensional search approach, because these estimators require access to L, which is assumed to be unknown in this work [15] [13]. We then report the performance of estimator (B.15) when there is a mismatch between the channel assumptions made for its derivation and the real channel conditions in which it is used. Specifically, we say that there is a mismatch if a LOS (NLOS) condition prevails in the channel, while the used estimator is the one derived under the assumption of NLOS (LOS). Otherwise there is a match. Remember that the factor q controls which of the LOS ($q = 1$) or NLOS ($q = 0$) condition holds. Table B.1 summarizes the settings used for the simulations of the considered OFDM system. Pilots with equal power are placed either with equal spacing (Uniform pilot pattern) or randomly (Random pilot pattern) in an OFDM symbol. A random pilot pattern is generated by sampling N_p pilots uniformly at random without replacement from $\mathbf{I}$. In the Monte Carlo simulation, we use the channel model in Section 2 Example 2.3.

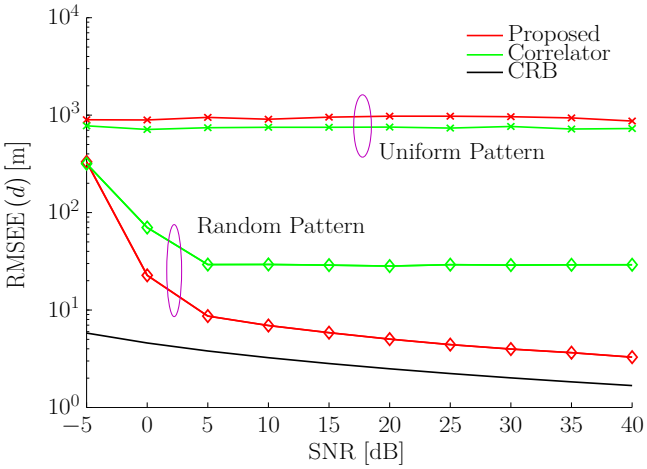


Fig. B.1: LOS scenario: RMSEE obtained by using different estimators and pilot patterns for a multi-path channel with RMS delay spread $\lambda = 50\,\text{ns}$ and average number of paths $\mu_L = 480$.

4.1 LOS Scenarios: Performance Evaluation Using Different Pilot Patterns and Estimators

Fig. B.1 shows the simulated root mean square estimation error (RMSEE) of d using estimator (B.15) and the correlator-based estimator. We observe that for both estimators, the uniform pilot pattern leads to outliers due to high sidelobes in the respective objective functions. This effect does not occur when the random pilot pattern is used. We observe in this case that the proposed estimator outperforms the correlator-based estimator. We then compare the results with the CRB [6], which is computed under the assumption that $\boldsymbol{\varepsilon}$ and $\mathbf{y}$ are jointly Gaussian. Since such assumption is not fulfilled here, the simulated RMSEE does not meet the CRB.

4.2 Performance Evaluation Under Different Channel Settings

From this point on, we only report the results obtained by employing a random pilot pattern. Fig. B.2 reports the simulated RMSEE in the LOS scenario. We observe that estimator (B.15) outperforms significantly the correlator-based estimator. As the average number of paths increases, the performance of the correlator-based estimator deteriorates. When μ_L is small, i.e. the Gaussian assumption is significantly violated, the RMSEE of estimator (B.15) noticeably deviates from the CRB. Such deviation becomes smaller as μ_L increases. Moreover, the accuracy of estimator (B.15) increases as the RMS delay spread

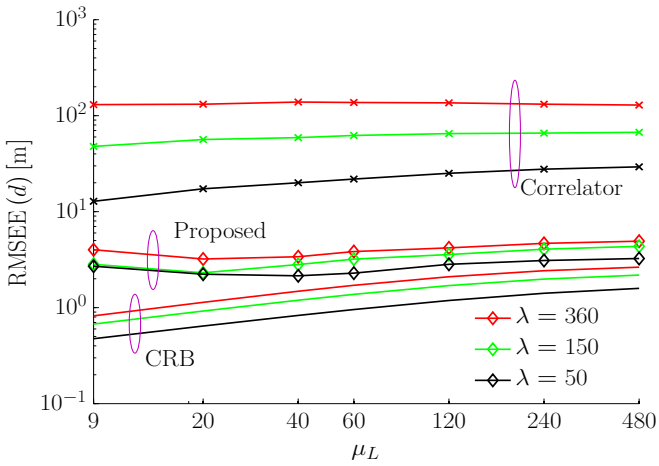


Fig. B.2: LOS scenario: RMSEE versus average number of paths μ_L for different values of the RMS delay spread λ at SNR = 40 dB.

decreases. Fig. B.3 depicts the cumulative distribution function (CDF) of the range errors in the LOS scenario. We observe that the medians are positive which indicates that positive errors are more frequent than negative errors. As μ_L decreases, the corresponding CDF shows a sharper slope and the median decreases accordingly. We remark that rare large outliers appear when μ_L is small, which lifts the overall RMSEE up as shown in Fig. B.2.

Fig. B.4 reports the simulated RMSEE in the NLOS scenario. Contrary to the LOS scenario, the RMSEE decreases as μ_L increases and the proposed estimator's performance becomes insensitive to the RMS delay spread when this parameter is large enough. When μ_L and the RMS delay spread of the channel are small, in which case the Gaussian assumption is significantly violated, both estimators yield large errors. However, compared to the correlator-based estimator, estimator (B.15) exhibits a promising performance gain when μ_L and the RMS delay spread of the channel are large. Fig. B.5 depicts the CDF of the range errors in NLOS scenarios. We notice that μ_L affects the proposed estimator's performance. Contrary to what was observed in Fig. B.3, the median increases as μ_L decreases. When $\mu_L = 9$, the median is at around 80 m, which explains the high RMSEE in Fig. B.4.

4.3 Performance Comparison in Conditions with Model Match and Mismatch

Fig. B.6 reports the simulated RMSEE when the assumptions used to derive estimator (B.15) match or mismatch the real channel propagation conditions.

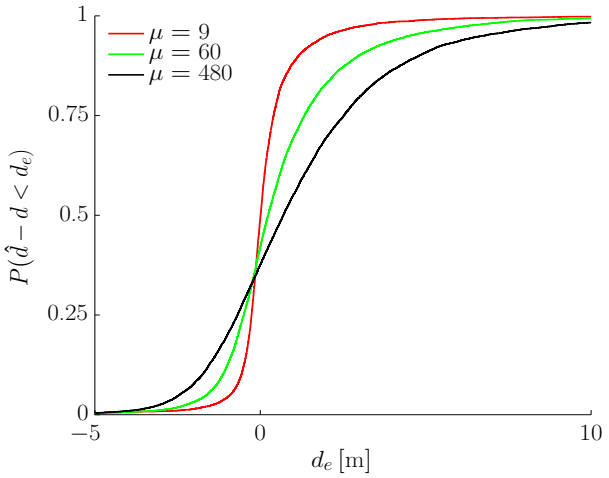


Fig. B.3: LOS scenario: Empirical CDF of the range errors obtained using (B.15) when RMS delay spread $\lambda = 360\,\text{ns}$ and SNR $= 40\,\text{dB}$.

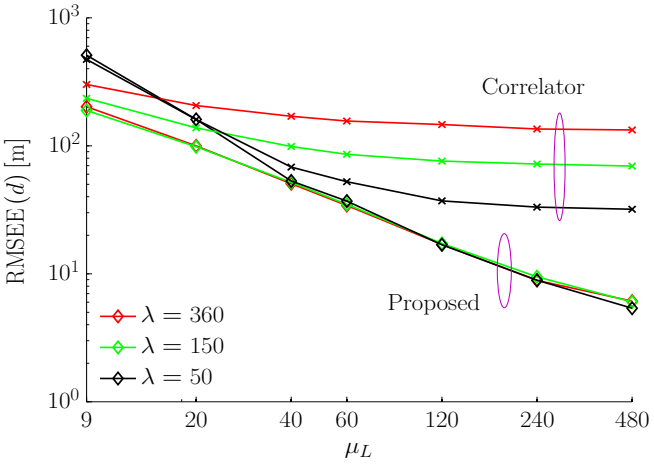


Fig. B.4: NLOS scenario: RMSEE versus the average number of paths μ_L for different values of RMS delay spread λ at SNR $= 40\,\text{dB}$.

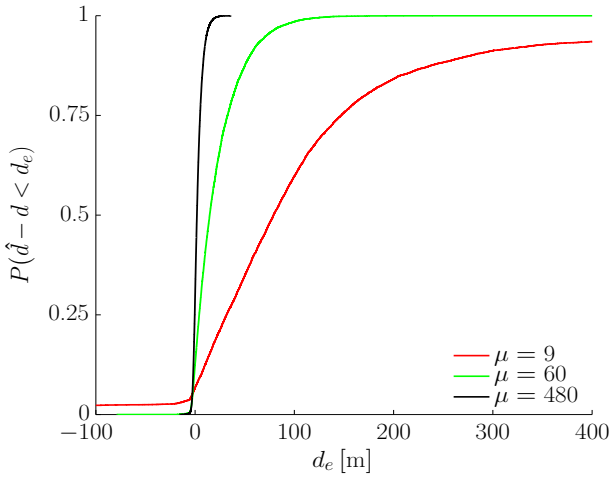


Fig. B.5: NLOS scenario: Empirical CDF of the range errors obtained using (B.15) when $\lambda = 360\,\text{ns}$ and $\text{SNR} = 40\,\text{dB}$.

Clearly, estimator (B.15) still outperforms the correlator-based estimator and both estimators benefit from the LOS propagation channel, which leads to lower RMSEEs. In case of a mismatch, estimator (B.15) performs worse than when there is a match. This is particularly noticeable in LOS conditions with small μ_L. Except for the matched case when the real channel is in LOS conditions, the RMSEE decreases as μ_L increases due to the Gaussian assumption becomes more realistic.

5 Conclusion

Using a channel model formulated as a point process, we demonstrate that the proposed approximate maximum likelihood estimator outperforms the correlator-based estimator. In the single-path scenario, the correlator-based estimator coincides with the proposed estimator. The proposed estimator does not require first-path detection, path separability, nor estimation of the number of path components. Though the invoked Gaussian assumptions is not fulfilled in typical channel conditions, the proposed estimator achieves promising range accuracy. An additional finding is that both proposed and correlator-based estimators achieve higher estimation accuracy when a random pilot pattern is employed rather than the uniform pilot pattern, as currently used in LTE.

Given the SNR and the covariance matrix of the channel, the complexity of the proposed estimator is tractable. The estimator accuracy depends on the RMS delay spread and the average number of path components. The proposed estimator achieves promising results in the LOS scenario even if there

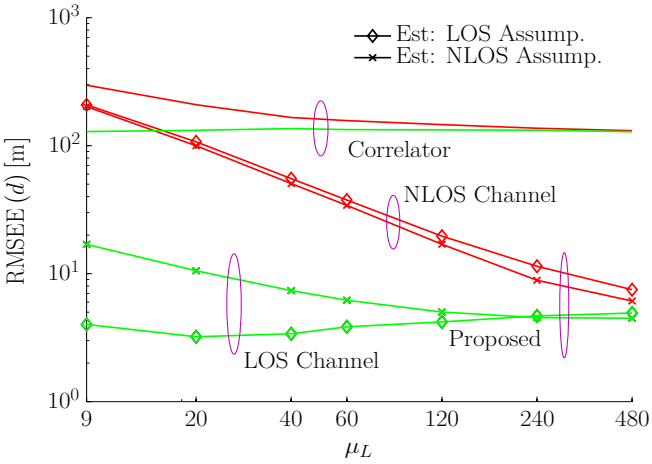


Fig. B.6: RMSEE versus average number of paths μ_L for RMS delay spread $\lambda = 360\,\text{ns}$ and SNR $= 40\,\text{dB}$. Green and red curves indicate the real channel conditions. "Est: LOS Assump." shows the RMMSE of estimator (B.15) assuming that the real propagation channel is in LOS conditions while "Est: NLOS Assump." denotes the RMMSE of estimator (B.15) assuming that the real channel is in NLOS conditions.

is a mismatch between the assumptions used to derive the proposed estimator and the real channel conditions. In the NLOS scenario, the average number of path components limits the estimator's performance in both matched and mismatched cases.

Acknowledgment

This work was supported by the EU FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306) and in part by the VIRTUOSO cooperative research project, funded by Intel Mobile Communications, Anite, Telenor, Aalborg University, and the Danish National Advanced Technology Foundation.

References

[1] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win. Ranging with ultrawide bandwidth signals in multi-path environments. *Proc. IEEE*, 97(2):404–426, 2009.

[2] F. Zhao, W. Yao, C. C. Logothetis, and Y. Song. Super-resolution TOA Estimation in OFDM Systems for Indoor Environments. In *IEEE International Conf. on Networking, Sensing and Control*, Apr. 2007.

[3] G. L. Turin, F. D. Clapp, T. L. Johnston, and S. B. Fine, and D. Lavry. A statistical model of urban multipath propagation. *IEEE Trans. Veh. Technol.*, 21(1):1–9, 1972.

[4] G. Steinbock, T. Pedersen, B. H. Fleury, W. Wang and R. Raulefs. Distance dependent model for the delay power spectrum of in-room radio channels. *IEEE Trans. Antennas Propag.*, 61(8):4327–4340, Aug. 2013.

[5] M. L. Jakobsen. *Modeling and Analysis of Stochastic Radio Channels.* PhD thesis, Aalborg University, 2014.

[6] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice-Hall PTR, 1998.

[7] J. Kingman. *Poisson Processes.* Oxford studies in probability. Clarendon Press, 1992.

[8] Y. Li, L. Cimini, and N. Sollenberger. Robust channel estimation for OFDM systems with rapid dispersive fading channels. *IEEE Trans. Wireless Commun.*, 46(7):902–915, 1998.

[9] M. L. Jakobsen, T. Pedersen, B. H. Fleury. Analysis of Stochastic Radio Channels with Temporal Birth-Death Dynamics: A Marked Spatial Point Process Perspective. *IEEE Trans. Antennas Propag.*, Apr. 2014.

[10] M. Triki and D. Slock. Mobile Localization for NLOS Propagation. In *Proc. IEEE 18th Int. Symposium on Personal, Indoor and Mobile Radio Commun., PIMRC*, pages 1–4, Sept. 2007.

[11] C. Mensing. *Location Determination in OFDM Based Mobile Radio Systems.* PhD thesis, Technische Universität München, 2013.

[12] O. Bialer, D. Raphaeli, and A. J. Weiss. Maximum-likelihood direct position estimation in dense multipath. *IEEE Trans. Veh. Technol.*, 62(5):2069–2079, June 2013.

[13] R. Adam and P. A. Hoeher. Semi-blind channel estimation for joint communication and positioning. *10th Workshop on Positioning Navigation and Commun.*, pages 1–5, 2013.

[14] A. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location. *IEEE Signal Process. Mag.*, 22(4):24–40, 2005.

[15] T. Wang, Y. Shen, S. Mazuelas, H. Shin and M. Z. Win. On OFDM ranging accuracy in multipath channels. *IEEE Systems Journal*, PP(99):1–11, 2013.

[16] D. Wang, M. Fattouche, and F. Ghannouchi. Fundamental limit of OFDM range estimation in a separable multipath environment. *Circuits, Systems, and Signal Processing*, 31(3), 2012.

Paper C

Bayesian Estimators for Direct Ranging in Multi-path Channels Using OFDM Pilot Signals

Lishuai Jing, Troels Pedersen, Bernard H. Fleury

The paper has been submitted to the
IEEE Transactions on Wireless Communications
Nov., 2014.

The layout has been revised.

Abstract

Direct ranging, which estimates the ranging parameter in one-step without estimating intermediate parameters, such as complex gains and delays, has recently been proposed. In this work, we address the problem of direct ranging using OFDM pilot signals via Bayesian estimators, namely the maximum-a-posteriori (MAP) estimator, the linear minimum mean-squared error (MMSE) estimator, and the MMSE estimator. However, these estimators cannot be directly applied to this problem since the probability density function of the received frequency-domain OFDM pilot signals is unknown. Instead, we propose approximate versions of them. Though the probability density function of the received signal is unknown, all its moments become readily available when Turin's classical multi-path channel model, which describes the path delays as a realization of a Poisson point process, is applied. We exploit this knowledge and derive a pth-order polynomial MMSE estimator which utilizes moments up to 2pth-order of the received signals for direct ranging. When $p = 1$, the proposed estimator coincides with the linear MMSE estimator, which is inapplicable to the addressed problem. To balance the computational complexity and the ranging accuracy, we choose to use the widely linear-quadratic MMSE estimator ($p = 2$). Monte Carlo simulations show that the proposed estimators achieve promising ranging accuracy gains and are fairly robust against model mismatches.

1 Introduction

Accurate localization is increasingly important for terrestrial wireless systems, in particular for OFDM systems such as WLAN, LTE and its extension LTE-A [20] [15]. One approach to improve the localization accuracy is to rely on high precision ranging techniques [4] [19] [21]. State-of-the-art ranging techniques (for instance the correlator-based estimator [15] [7], the energy-detection-based estimator [4], and the multidimensional-search-based estimator [21] [18]) follow a two-step approach. First, parameters, such as the received signal strength, the time of arrival, the bias induced by none-line-of-sight (NLOS) propagation, etc., are estimated from the received signal. Then, these estimates are used for ranging. The correlator-based and energy-detection-based estimators are easy to implement, but both suffer from inferior ranging accuracy when the system bandwidth is not sufficiently high or the line-of-sight (LOS) component is not strong enough [4]. The multidimensional-search-based estimator proposed in [21], on the other hand, offers better ranging accuracy when the number of paths in the channel response is known in advance. But, in general, estimating the exact number of paths is a difficult task. Even if this number is known, the complexity of performing a multidimensional search is high. Potentially, lower

complexity iterative schemes, such as the SAGE (space-alternating generalized expectation-maximization) algorithm [2], can be applied. These schemes, however, require the knowledge of the number of path components and might converge to a local maximum.

An alternative approach is direct ranging, which estimates the range in one step. In [14], Jing *et al.* employ a model of the channel impulse response that formulates the random set of path delays as a realization of a point process. When the number of path components in the channel response and the separability condition of these components are unknown and the delay power spectrum exhibits an exponential decay, the vector of received OFDM pilot signals is not Gaussian. In such scenarios, the authors propose an approximate maximum likelihood estimator (referred to as GAML estimator) that uses a Gaussian approximation of the distribution of the received pilot vector to increase the ranging accuracy. This estimator does not rely on first-path detection, any separability condition, and the knowledge of the number of path components. The GAML estimator assumes that the SNR and the delay power spectrum are known. It has a tractable complexity and offers significantly higher ranging accuracy than the non-coherent correlator-based estimator. However, as a consequence of the Gaussian approximation, the GAML estimator only utilizes the first- and second-order statistics of the received signal. In case higher-order moments of the received signal and the prior distribution of the range are accessible, this knowledge can potentially be exploited for improving ranging accuracy. An overview on utilizing the higher-order statistics and prior knowledge of the parameters of interest in signal processing can be found in [9].

In the present contribution, we address the problem of direct ranging using OFDM pilot signals in multi-path channels via Bayesian estimators. Inspired by the Gaussian approximation used in [14], we derive an approximate maximum-a-posteriori (MAP) estimator and an approximate minimum-mean-square-error (MMSE) estimator. Compared to [14], we impose a prior distribution for the range parameter, which makes the application of Bayesian inference possible. In addition, we show that when the classical Turin's channel model [5] is applied, all moments of the frequency-domain channel vector and therefore of the received OFDM pilot signal vector become computable, while their joint probability density functions (pdfs) are not. Having access to these higher-order moments, we propose to use a pth-order polynomial MMSE estimator for direct ranging. Such estimators have been applied for the estimation of signal amplitudes using higher-order statistics [16]. A pth-order polynomial MMSE estimator requires the moments up to order $2p$ of the received signal. In contrast to many other engineering problems [16] in which the required higher-order statistics need to be estimated to employ the pth-order polynomial MMSE estimator, in our particular application context, we can compute them in a closed form expression. The pth-order polynomial MMSE estimator requires the signal-to-noise-ratio (SNR). But it is unbiased and inherits all

merits of the GAML estimator: it avoids first-path detection problem and the requirement of knowing the exact number of path components and it relaxes the separability condition. When $p = 1$, the pth-order polynomial MMSE estimator coincides with the linear MMSE estimator. However, this estimator collapses to the mean of the prior due to the invoked assumptions on the multi-path channel model. Therefore it is inapplicable to the problem at hand. The widely linear-quadratic MMSE estimator, i.e. $p = 2$, can be applied, however. We demonstrate by means of Monte Carlo simulation that the widely linear-quadratic MMSE, the approximate MAP, and the approximate MMSE estimators significantly outperform the correlator-based estimator.

2 System and Signal Model

We consider a single-input single-output OFDM setup with N pilots indexed by $n = 1, 2, \ldots, N$ in one OFDM symbol of time duration T. An OFDM symbol is generated by multiplexing a sequence of data symbols and known pilot symbols onto a number of orthogonal sub-carriers. The adjacent sub-carrier spacing is $\Delta f = \frac{1}{T}$. A cyclic prefix of duration T_{cp} is appended to prevent inter-symbol and inter-carrier interference.

We address the problem of the estimation of the range parameter r based on the received pilot signals. In contrast to [14], we consider here r to be a random variable having a-priori pdf $p(r)$ with mean μ_r and variance σ_r^2. The multi-path channel is assumed to be time invariant during the transmission of one OFDM symbol. Removing the cyclic prefix and concatenating the received pilot signals in the observation vector $\mathbf{y}$, we obtain the signal model in the frequency domain [14]:

$$\mathbf{y} = \mathbf{A}\big[\boldsymbol{\varphi}(r) \odot \underbrace{(q\alpha_0\mathbf{1} + \boldsymbol{\varepsilon})}_{\tilde{\boldsymbol{\varepsilon}}}\big] + \mathbf{n}. \tag{C.1}$$

In (C.1), $\mathbf{A} = \text{diag}\{a_1, \ldots, a_N\}$ is a diagonal matrix with a_i, $i = 1, \ldots, N$, denoting the ith pilot symbol, $\odot$ is the Hadamard product, $\mathbf{1}$ is an all-ones vector, and $\mathbf{n}$ is a white circular (in the sense defined in [17]) complex Gaussian noise vector with component variance σ^2. The column vector containing the ranging parameter reads

$$\boldsymbol{\varphi}(r) = [\varphi_1, \ldots, \varphi_N]^T,$$

where

$$\varphi_n = e^{-\jmath 2\pi f_n \frac{r}{c}}$$

with f_n denoting the frequency of pilot n, c being the speed of light, and $\jmath = \sqrt{-1}$. As suggested by Jing *et al.* [14], we factorize the channel frequency response in a range-dependent term $\boldsymbol{\varphi}(r)$ and a range-independent term $\tilde{\boldsymbol{\varepsilon}}$ for

direct ranging purpose. The simplifying assumption that $\boldsymbol{\varepsilon}$ is independent of r may or may not be realistic. We leave the investigation of more sophisticated distance-dependent channel models, such as the one presented in [6], to future works.

The indicator parameter $q \geq 0$ specifies the setting of the LOS path component. In a LOS condition i.e. $q > 0$, q adjusts the power of the LOS component. When $q = 0$, the system operates in a NLOS condition. The mth entry of the column vector $\boldsymbol{\varepsilon}$ reads

$$\varepsilon_m = \sum_{l=1}^{L} \alpha_l e^{-j2\pi f_m \tau_l}, \tag{C.2}$$

where α_l is the complex gain and τ_l is the excess delay of path l with respect to the delay $\frac{r}{c}$ induced by the geometric distance between the transmitter and receiver. As proposed by Turin [5], we model the random set of excess delays as a Poisson point process $\mathcal{T} = \{\tau_1, \tau_2, \ldots\}$ with intensity function $\rho(\tau)$ which gives the average number of points in $\mathcal{T}$ per time unit. Under these assumptions, the elements in $\mathcal{T}$ and the number of multi-path components $L = |\mathcal{T}|$ are random. In addition, depending on the shape of $\rho(\tau)$, L may not be finite. We further assume that the complex gains are circular random variables and conditionally independent. Thus, they have mean zero and conditional second-order moments

$$E[\alpha_l \alpha_{l'}^* | \tau_l, \tau_{l'}] = \begin{cases} \sigma_\alpha^2(\tau_l), & l = l' \\ 0, & \text{otherwise.} \end{cases} \tag{C.3}$$

With these assumptions, it is readily verified that $\mathbf{y}$ is circular. The delay power spectrum is given by [11]

$$P(\tau) = \sigma_\alpha^2(\tau)(\rho(\tau) + q^2\delta(\tau))$$

with δ denoting the Dirac delta function.

3 Approximate MAP and Approximate MMSE Direct Ranging Techniques

When the prior $p(r)$ is known, we can use Bayesian inference to estimate r. Well-known Bayesian estimators include the maximum a posteriori (MAP) estimator, the minimum mean-squared error (MMSE) estimator, and the linear MMSE estimator. When higher-order moments of the observation vector $\mathbf{y}$ are available, polynomial MMSE estimators [16] can be applied. We remark that all estimators discussed next are "direct" range estimators since no intermediate parameters, such as delays, complex gains, etc., are estimated [14].

3.1 Approximate MAP Estimator

The MAP estimator selects a value $\hat{r}$ that maximizes the posterior probability density function $p(r|\mathbf{y})$ [12]:

$$\begin{aligned} \hat{r}_{\mathrm{MAP}}(\mathbf{y}) &= \arg\max_r p(r|\mathbf{y}) \\ &= \arg\max_r p(\mathbf{y}|r)p(r). \end{aligned} \tag{C.4}$$

In the second line, $p(\mathbf{y}|r)$ denotes the likelihood function of r given the observation $\mathbf{y}$. Due to the invoked assumptions on the channel model, $p(\mathbf{y}|r)$ is unknown. If the number of path components is assumed to be known as in [21], $p(\mathbf{y}|r)$ may become computable. But accurately estimating the number of path components can be difficult given the system bandwidth limitation.

Here, we follow the alternative approach of approximating $p(\mathbf{y}|r)$ with a Gaussian pdf $\tilde{p}(\mathbf{y}|r)$ with the same first- and second-order moments [14], i.e. $\tilde{p}(\mathbf{y}|r)$ is a second-order maximum entropy approximation of $p(\mathbf{y}|r)$ [10]. This approximation is exact if $\tilde{\boldsymbol{\varepsilon}}$ is a Gaussian random vector, which is a reasonable approximation when L is large and $\sigma_\alpha^2(\tau)$ is a constant. However, in more realistic propagation conditions, e.g. when the delay power spectrum decays exponentially, the Gaussian approximation can be inaccurate. The Gaussian approximation leads to an estimator that can be derived analytically. Using $\tilde{p}(\mathbf{y}|r)$ instead of $p(\mathbf{y}|r)$ in (C.4) yields

$$\hat{r}_{\mathrm{AMAP}}(\mathbf{y}) = \arg\max_r \ln \tilde{p}(\mathbf{y}|r) + \ln p(r), \tag{C.5}$$

where

$$\tilde{p}(\mathbf{y}|r) = \frac{1}{\pi^N \det(\boldsymbol{C}_{\mathbf{y}}(r))} e^{-\mathbf{y}^H \boldsymbol{C}_{\mathbf{y}}^{-1}(r)\mathbf{y}} \tag{C.6}$$

with $\det(\cdot)$ denoting the determinant. The conditional covariance matrix $\mathbf{C}_{\mathbf{y}}(r) = E[\mathbf{y}\mathbf{y}^H|r]$ reads

$$\mathbf{C}_{\mathbf{y}}(r) = \mathbf{A}\boldsymbol{\Phi}(r)\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}\boldsymbol{\Phi}^H(r)\mathbf{A}^H + \sigma^2\mathbb{I} \tag{C.7}$$

with $\boldsymbol{\Phi}(r) = \mathrm{diag}(\boldsymbol{\varphi}(r))$, $(\cdot)^H$ denoting conjugate transposition, $\mathbb{I}$ being the identity matrix, and $\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}} = E[\tilde{\boldsymbol{\varepsilon}}\tilde{\boldsymbol{\varepsilon}}^H]$. The (m,n)th entry of $\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}$ reads [11]

$$[\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}]_{mn} = \mathcal{F}\{P(\tau)\}(f_m - f_n), \tag{C.8}$$

where $\mathcal{F}$ denotes the Fourier transform. Using the eigenvalue decomposition $\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, we can recast (C.7) as

$$\mathbf{C}_{\mathbf{y}}(r) = \boldsymbol{\Phi}(r)\mathbf{G}\boldsymbol{\Phi}^H(r)$$

with

$$\mathbf{G} = \mathbf{A}\mathbf{U}(\boldsymbol{\Lambda} + \mathbf{I}\sigma^2/E_s)\mathbf{U}^H\mathbf{A}^H,$$

where $E_s = E[|a_n|^2]$ is the pilot symbol power. Since $\mathbf{\Phi}(r)$ is unitary, $\det(\mathbf{C_y}(r)) = \det(\mathbf{G})$, i.e. the determinant of $\mathbf{C_y}(r)$ does not depend on r. Thus, the log-likelihood $\ln \tilde{p}(\mathbf{y}|r)$ is of the form

$$\ln \tilde{p}(\mathbf{y}|r) = -\mathbf{y}^H \mathbf{\Phi}^H(r)\mathbf{G}^{-1}\mathbf{\Phi}(r)\mathbf{y} + \text{constant} \tag{C.9}$$

with

$$\mathbf{G}^{-1} = \mathbf{AU}(\mathbf{\Lambda} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}\mathbf{U}^H\mathbf{A}^H. \tag{C.10}$$

Thus, the approximate MAP estimator reads

$$\hat{r}_{\text{AMAP}}(\mathbf{y}) = \arg\max_r -\mathbf{y}^H \mathbf{\Phi}(r)\mathbf{G}^{-1}\mathbf{\Phi}^H(r)\mathbf{y} + \ln p(r). \tag{C.11}$$

Since the matrices $\mathbf{AU}$ and $\mathbf{\Lambda}$ can be pre-computed and stored, the inversion of $\mathbf{G}$ amounts to compute the diagonal matrix $(\mathbf{\Lambda} + \frac{\sigma^2}{E_s}\mathbf{I})^{-1}$. This circumvents the brute force inversion of $\mathbf{C_y}(r)$ in (C.6) and thereby reduces the complexity of the estimator. We remark that the matrices $\mathbf{U}$ and $\mathbf{\Lambda}$ depend on $P(\tau)$. Therefore, they should be precomputed and stored for each setting of the parameters of $P(\tau)$.

The approximate MAP estimator is readily related to the GAML estimator proposed in [14]. When r is uniformly distributed on an interval $\mathcal{D}$, (C.11) simplifies to

$$\hat{r}_{\text{AMAP}}(\mathbf{y}) = \arg\min_{r\in\mathcal{D}} \mathbf{y}^H \mathbf{\Phi}(r)\mathbf{G}^{-1}\mathbf{\Phi}^H(r)\mathbf{y}. \tag{C.12}$$

The approximate MAP estimator (C.12) differs from the GAML estimator [14] in the restriction to $\mathcal{D}$ of the argmin operation. It can be shown that, when r is a deterministic parameter, the non-coherent correlator-based estimator [15] [4] [22] is a limiting case of (C.12) as $q \to \infty$, which corresponds to a single-path channel.

3.2 Approximate MMSE Estimator

The standard MMSE estimator for the problem at hand reads

$$\begin{aligned} \hat{r}_{\text{MMSE}}(\mathbf{y}) &= E[r|\mathbf{y}] \\ &= \int r p(r|\mathbf{y})\mathrm{d}r \\ &= \frac{1}{\int p(\mathbf{y}|r)p(r)\mathrm{d}r}\int r p(\mathbf{y}|r)p(r)\mathrm{d}r, \end{aligned} \tag{C.13}$$

where the last step follows by invoking Bayes' rule. Using the same line of arguments as in Section 3.1, we replace $p(\mathbf{y}|r)$ with $\tilde{p}(\mathbf{y}|r)$ in (C.13) to obtain the approximate MMSE estimator

$$\hat{r}_{\text{AMMSE}}(\mathbf{y}) = \frac{1}{\int \tilde{p}(\mathbf{y}|r)p(r)\mathrm{d}r}\int r\tilde{p}(\mathbf{y}|r)p(r)\mathrm{d}r. \tag{C.14}$$

Since the terms $\det(\mathbf{C}_{\mathbf{y}}(r)) = \det(\mathbf{G})$ in the numerator and denominator cancel each other out, (C.14) takes the form

$$\hat{r}_{\mathrm{AMMSE}}(\mathbf{y}) = \frac{1}{\int p(r) e^{-\mathbf{y}^H \boldsymbol{C}_{\mathbf{y}}^{-1}(r)\mathbf{y}} \mathrm{d}r} \int r e^{-\mathbf{y}^H \boldsymbol{C}_{\mathbf{y}}^{-1}(r)\mathbf{y}} p(r) \mathrm{d}r. \tag{C.15}$$

For a general $p(r)$, a closed-form expression of the right hand side of (C.14) is difficult to obtain or does not exist [1]. Inversion of the conditional covariance matrix $\mathbf{C}_{\mathbf{y}}(r)$ can be performed efficiently as shown before, see (C.9) and (C.10). As a result, the complexity of implementing (C.15) amounts to that of the numerical evaluation of two 1D integrals, which can be done fairly accurately via many existing numerical integration methods.

Compared to the GAML estimator and the approximate MAP estimator (C.11), the approximate MMSE estimator has the advantage that no search procedure is needed. However, it exhibits the same drawback as the approximate MAP estimator: its performance is expected to deteriorate when the average number of path components in (C.2) is small or the delay power spectrum exhibits a fast exponential decay. In this case, the assumption that $\mathbf{y}$ is jointly Gaussian may not be well justified since the central limit theorem does not apply.

4 A pth-order Polynomial Direct Ranging Estimator

When the MMSE estimator cannot be obtained in closed form and its implementation becomes too complex, linear estimators come into use. Since the observation $\mathbf{y}$ is complex, widely linear estimators can be applied [17]. As shown in Appendix B, for the addressed ranging problem the linear and widely linear MMSE estimators, however, are inapplicable since $\boldsymbol{C}_{r\mathbf{y}} = E[(r - \mu_{\mathrm{r}})\mathbf{y}^H] = \mathbf{0}$, as a result of $\boldsymbol{\varepsilon}$ and r being independent.

We show in Section 4.1 that, with our selection of the stochastic model of the radio channel, it is possible to compute analytically the moments of any order of the observation $\mathbf{y}$. We exploit this fact and propose estimators of the range that are polynomial functions of $\mathbf{y}$. Specifically, for a given $p = 1, 2, \ldots$, a pth-order polynomial estimator of r is given by

$$\begin{aligned} \hat{r}(\mathbf{y}) \quad = \quad & \beta_0^* + \sum_{i_1^{\Diamond}} \beta_{i_1^{\Diamond}}^* y_{i_1}^{\Diamond_{i_1}} + \cdots \\ & + \sum_{i_1^{\Diamond},\ldots,i_p^{\Diamond}} \beta_{i_1^{\Diamond},\ldots,i_p^{\Diamond}}^* y_{i_1}^{\Diamond_{i_1}} \cdots y_{i_p}^{\Diamond_{i_p}}, \end{aligned} \tag{C.16}$$

where we adopt the notation $y_i^{\Diamond_i}$ from [17] with $\Diamond_i$ indicating whether or not y_i is conjugated and $i^{\Diamond} = (i, \Diamond_i)$. Thus, the sum over $i^{\Diamond}$ includes $2N$ terms.

We recast the expression (C.16) in a vector form as

$$\hat{r}(\mathbf{y}) \quad = \quad \beta_0^* + \boldsymbol{\beta}^H \boldsymbol{z}, \tag{C.17}$$

where the column vector $\boldsymbol{z}$ has entries $y_{i_1}^{\lozenge_{i_1}} \cdots y_{i_p}^{\lozenge_{i_p}}$ with the associated coefficients arranged in $\boldsymbol{\beta}$. Note that the subsequent results hold irrespective of the arrangement of the entries in $\boldsymbol{z}$. For $p = 1$, (C.17) is the so-called widely linear MMSE estimator[1] [17] with $\boldsymbol{z}$ defined as $\boldsymbol{z} = [\mathbf{y}^T \mathbf{y}^H]^T$. For $p = 2$, (C.17) is referred to as the widely linear MMSE estimator [17] with

$$\boldsymbol{z} = \begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \\ \mathbf{y} \otimes \mathbf{y} \\ \mathbf{y}^* \otimes \mathbf{y}^* \\ \mathbf{y} \otimes \mathbf{y}^* \end{bmatrix}, \tag{C.18}$$

where $\otimes$ denotes the Kronecker product.

The coefficients of the optimal pth-order polynomial estimator minimize the mean-square-error $E[|r - \hat{r}(\mathbf{y})|^2]$. By invoking the orthogonality principle, these coefficients are the solutions of the set of equations

$$E[r - \hat{r}(\mathbf{y})] = 0, \tag{C.19}$$

$$E[(r - \hat{r}(\mathbf{y}))\boldsymbol{z}^H] = \mathbf{0}. \tag{C.20}$$

Solving (C.19) and (C.20) yields

$$\beta_0^* = \mu_r - \boldsymbol{\beta}^H E[\boldsymbol{z}], \quad \boldsymbol{\beta}^H = \boldsymbol{C}_{r\boldsymbol{z}} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1}, \tag{C.21}$$

where

$$\boldsymbol{C}_{r\boldsymbol{z}} = E[(r - \mu_r)(\boldsymbol{z} - E[\boldsymbol{z}])^H],$$
$$\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}} = E[(\boldsymbol{z} - E[\boldsymbol{z}])(\boldsymbol{z} - E[\boldsymbol{z}])^H].$$

Consequently, the pth-order polynomial MMSE estimator of r reads

$$\hat{r}(\mathbf{y}) = \mu_r + \boldsymbol{C}_{r\boldsymbol{z}} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1} (\boldsymbol{z} - E[\boldsymbol{z}]). \tag{C.22}$$

Accordingly, the root-mean-square-error (RMSE) of the estimator is given by

$$\sqrt{E[|r - \hat{r}(\mathbf{y})|^2]} = \sqrt{\sigma_r^2 - \boldsymbol{C}_{r\boldsymbol{z}} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1} \boldsymbol{C}_{r\boldsymbol{z}}^H}. \tag{C.23}$$

The estimator (C.22) requires the moments up to order $2p$ of $\mathbf{y}$ to be known. By adopting Turin's channel model, we can in fact compute all its moments of any order. This is a remarkable benefit of the mechanism generating the path

[1] Strictly speaking, the pth-order MMSE estimator (C.17) is not a linear estimator. We stick to the terminology used in [17].

delays using a Poisson point process. Without this feature, we would need to estimate the required higher-order moments, like many other engineering problems do [16]. In addition, as compared to the approximate MAP estimator (C.11) and the approximate MMSE estimator (C.14), another advantage of the pth-order polynomial MMSE estimator is that no approximation of the pdf of $\mathbf{y}$ is needed as long as all its moments up to order $2p$ are known.

In writing (C.22), we assume that $\boldsymbol{C}_{\boldsymbol{zz}}$ is invertible. In principle, as p increases, a higher ranging accuracy can be achieved. But the complexity of computing $\boldsymbol{C}_{\boldsymbol{zz}}^{-1}$ grows exponentially as p increases. When $p = 2$, computing $\boldsymbol{C}_{\boldsymbol{zz}}^{-1}$ with dimension $(2N + 3N^2) \times (2N + 3N^2)$ can be time consuming for large N. As shown in Appendix A, for circular $\mathbf{y}$, $\boldsymbol{C}_{\boldsymbol{zz}}$ is a block diagonal matrix. This structure reduces the complexity of computing $\boldsymbol{C}_{\boldsymbol{zz}}^{-1}$. To further decrease the computational effort without increasing (C.23) too much, one might need to judiciously choose the terms of the sum in (C.17). Hence the RMSE in (C.23) may be used as a criterion for this selection.

4.1 Higher-order Cumulants and Moments

Application of the pth-order polynomial MMSE estimator in (C.22) necessitates computation of the two covariance matrices $\boldsymbol{C}_{r\boldsymbol{z}}$ and $\boldsymbol{C}_{\boldsymbol{zz}}$. By the law of total expectation, we have

$$\boldsymbol{C}_{r\boldsymbol{z}} = E\Big[rE[\boldsymbol{z}^H|r]\Big] - \mu_r E\Big[E[\boldsymbol{z}|r]\Big], \tag{C.24}$$

$$\boldsymbol{C}_{\boldsymbol{zz}} = E\Big[E[\boldsymbol{z}\boldsymbol{z}^H|r]\Big] - \Big|E\Big[E[\boldsymbol{z}|r]\Big]\Big|^2. \tag{C.25}$$

Thus, computation of the covariance matrices using (C.24) and (C.25) requires access to the conditional moments of the entries of $\mathbf{y}$ up to order $2p$. To do so, we follow the approach of first calculating cumulants and then use these to obtain the required moments.

We define the cumulant generator (cumulant generating function) of a complex random vector $\boldsymbol{\nu} = [\nu_1, \ldots, \nu_N]$ as

$$\psi_{\boldsymbol{\nu}}(\boldsymbol{w}) = \ln E[e^{\jmath\Re(\boldsymbol{w}^H\boldsymbol{\nu})}],$$

where $\Re$ denotes real part. Let $\boldsymbol{\nu}_{\mathbf{I}} = (\nu_{i_1}^{\diamondsuit_{i_1}}, \cdots, \nu_{i_k}^{\diamondsuit_{i_k}})$ denote a selection of $k = k_1 + k_2$ entries of $\boldsymbol{\nu}$ of which k_2 are conjugated. The corresponding cumulant is defined as

$$\kappa(\boldsymbol{\nu}_{\mathbf{I}}) = \Big(\frac{2}{\jmath}\Big)^k \frac{\partial^k \psi_{\boldsymbol{\nu}}(\boldsymbol{w})}{\partial w_{\mathbf{I}}}\bigg|_{\boldsymbol{w}=\mathbf{0}} \tag{C.26}$$

with $\partial w_{\mathbf{I}} = (\partial w_{i1}^{\diamondsuit_{i_1}})^* \cdots (\partial w_{i_k}^{\diamondsuit_{i_k}})^*$. The conditional moments and cumulants

are related via [9] [8]

$$\mathrm{Mom}(\boldsymbol{\nu}_{\mathbf{I}}) = E[\nu_{i_1}^{\Diamond_{i_1}} \cdots \nu_{i_k}^{\Diamond_{i_k}}] = \sum_{\pi} \prod_{\mathbf{b}\in\pi} \kappa(\mathbf{b}) \tag{C.27}$$

where π runs through all partitions of the sequence $\boldsymbol{\nu}_{\mathbf{I}}$ and $\mathbf{b}$ runs through all sets of the partition π. In the case of a circular vector $\boldsymbol{\nu}$, the task of evaluating (C.27) is alleviated by noticing that all cumulants, but those corresponding to $k_1 = k_2$, are zero. To compute the conditional cumulants, we replace $\nu_{\mathbf{I}}$ by $\nu_{\mathbf{I}}|r$ in the above definitions.

We now compute the conditional cumulants of $\mathbf{y}$ by applying Campbell's theorem. From (B.1), given r, the cumulants of $\mathbf{y}$ is a sum of three terms:

$$\kappa(\mathbf{y}_{\mathbf{I}}|r) = \kappa(\mathbf{n}_{\mathbf{I}}) + \kappa((q\alpha_0\mathbf{A}\boldsymbol{\varphi}(r))_{\mathbf{I}}|r) + \kappa((\mathbf{A}\boldsymbol{\varphi}(r)\odot\boldsymbol{\varepsilon})_{\mathbf{I}}|r). \tag{C.28}$$

These terms are the cumulants of the noise, the signal received via a LOS path, and the remaining multi-path components, respectively.

All cumulants of the circular Gaussian noise vector $\mathbf{n}$ vanish except the second-order cumulants, which are of the form

$$\kappa(n_i, n_i^*) = \sigma^2, \quad i = 1, \ldots, N.$$

The second term in (C.28) simplifies due to the homogeneity property of cumulants:

$$\kappa\Big((q\alpha_0\mathbf{A}\boldsymbol{\varphi}(r))_{\mathbf{I}}\Big|r\Big) = \kappa\Big(\alpha_0^{\Diamond_{i_1}}, \ldots, \alpha_0^{\Diamond_{i_k}}\Big) q^k \prod_{i'=i_1,\ldots,i_k} a_{i'}^{\Diamond_{i'}} \varphi_{i'}^{\Diamond_{i'}}. \tag{C.29}$$

Since α_0 is circular Gaussian, we have

$$\kappa\Big(\alpha_0^{\Diamond_{i_1}}, \ldots, \alpha_0^{\Diamond_{i_k}}\Big) = \begin{cases} \sigma_\alpha^2(0) & \text{for } k_1 = k_2 = 1 \\ 0 & \text{otherwise.} \end{cases}$$

As expected, these cumulants all vanish in a NLOS condition, i.e. when $q = 0$.

Invoking homogeneity property and using the assumption that $\boldsymbol{\varepsilon}$ is independent of r, the third term in (C.28) becomes

$$\kappa\Big((\mathbf{A}\boldsymbol{\varphi}(r)\odot\boldsymbol{\varepsilon})_{\mathbf{I}}\Big|r\Big) = \kappa(\boldsymbol{\varepsilon}_{\mathbf{I}}) \prod_{i'=i_1,\ldots,i_k} a_{i'}^{\Diamond_{i'}} \varphi_{i'}^{\Diamond_{i'}}. \tag{C.30}$$

The cumulant $\kappa(\boldsymbol{\varepsilon}_{\mathbf{I}})$ is computed from the cumulant generator $\psi_{\boldsymbol{\varepsilon}}(\boldsymbol{w})$, which can be derived via the Campbell's theorem [13] as

$$\psi_{\boldsymbol{\varepsilon}}(\boldsymbol{w}) = \int (\Psi_{\boldsymbol{\xi}|\tau}(\boldsymbol{w}) - 1)\rho(\tau)\mathrm{d}\tau. \tag{C.31}$$

Here, $\boldsymbol{\xi}|\tau$ denotes a circular complex Gaussian vector with entries $\xi_i = \alpha_\tau e^{-\jmath 2\pi f_i \tau}$, $i = 1, 2, \ldots, N$, and conditional characteristic function

$$\begin{aligned} \Psi_{\boldsymbol{\xi}|\tau}(\boldsymbol{w}) &= E[\exp(\jmath \Re(\boldsymbol{w}^H \boldsymbol{\xi}))|\tau] \\ &= e^{-\frac{1}{4}\boldsymbol{w}^H \boldsymbol{C}_{\boldsymbol{\xi}|\tau} \boldsymbol{w}}, \end{aligned} \tag{C.32}$$

where $\boldsymbol{C}_{\boldsymbol{\xi}|\tau} = E[\boldsymbol{\xi}\boldsymbol{\xi}^H|\tau] = \sigma_\alpha^2(\tau)\boldsymbol{\varphi}(c\tau)\boldsymbol{\varphi}^H(c\tau)$. The cumulants of $\boldsymbol{\varepsilon}$ can be evaluated by applying (C.31) to (C.26). Doing so, we observe that the moments of $\boldsymbol{\xi}|\tau$ need to be computed. By inspection of (C.27) and using the fact that $\boldsymbol{\xi}|\tau$ is Gaussian with zero-mean yield

$$E[\xi_{i_1}^{\diamondsuit_{i_1}} \ldots \xi_{i_k}^{\diamondsuit_{i_k}}|\tau] = k_1! e^{-\jmath 2\pi f_{\text{dif}}\tau} \sigma_\alpha^{2k_1}(\tau) \delta_{k_1,k_2},$$

where δ_{k_1,k_2} denotes the Kronecker delta and $f_{\text{dif}} = (\pm f_{i_1} \pm f_{i_2} \cdots \pm f_{i_k})$ with the ith term negated if $\diamondsuit_i$ means complex conjugation, and positive sign otherwise. Therefore,

$$\kappa(\boldsymbol{\varepsilon}_{\mathbf{I}}) = k_1! \int e^{-\jmath 2\pi f_{\text{dif}}\tau} \sigma_\alpha^{2k_1}(\tau) \rho(\tau) \mathrm{d}\tau. \tag{C.33}$$

The conditional cumulants of $\mathbf{y}$ can now be computed by insertion into (C.28). Thereafter, the conditional moments are obtained from (C.27), which are inserted in (C.24) and (C.25) to compute the (unconditional) moments of $\mathbf{y}$.

We remark that due to the particular form of the cumulants in (C.29) and (C.30), computation of the expectations over r in (C.24) and (C.25) amount to computing the factors

$$\begin{aligned} E\Big[\prod_{i'=i_1,\ldots,i_k} \varphi_{i'}^{\diamondsuit_{i'}}\Big] &= \int p(r) e^{-\jmath 2\pi f_{\text{dif}} r/c} \mathrm{d}r \\ &= \mathcal{F}\{p(r)\}(f_{\text{dif}}/c) \end{aligned} \tag{C.34}$$

$$\begin{aligned} E\Big[r \prod_{i'=i_1,\ldots,i_k} \varphi_{i'}^{\diamondsuit_{i'}}\Big] &= \int r p(r) e^{-\jmath 2\pi f_{\text{dif}} r/c} \mathrm{d}r \\ &= \mathcal{F}\{rp(r)\}(f_{\text{dif}}/c). \end{aligned} \tag{C.35}$$

We remark that (C.34) is the value of the characteristic function of the prior pdf $p(r)$ evaluated at $2\pi f_{\text{dif}}/c$. Upon evaluating these expressions analytically or numerically, we can compute all required higher-order moments of $\mathbf{y}$ for a particular prior $p(r)$.

4.2 Widely Linear Quadratic MMSE Estimator

We now illustrate a specific example of the pth-order polynomial MMSE estimator: the widely linear-quadratic MMSE estimator. Its general form for a

complex signal model reads

$$\begin{aligned}\hat{r}_{LQ}(\mathbf{y}) &= \beta_0^* + \sum_i g_i^* y_i + \sum_i h_i^* y_i^* \\ &\quad + \sum_{i,j} \gamma_{i,j}^* y_i y_j + \sum_{i,j} \vartheta_{i,j}^* y_i^* y_j^* + \sum_{i,j} \beta_{i,j}^* y_i y_j^*. \end{aligned} \tag{C.36}$$

Since $\boldsymbol{C}_{r\mathbf{y}} = \mathbf{0}$, it follows that $g_i^* = h_i^* = 0$, $i = 1, \ldots, N$. It is shown in Appendix C that since $E[y_i y_j]$, $E[y_i^* y_j^*]$, $E[r y_i y_j]$, and $E[r y_i^* y_j^*]$, $i, j = 1, \ldots, N$, are zero, the corresponding filter coefficients $\gamma_{i,j}^*$ and $\vartheta_{i,j}^*$, $i, j = 1, \ldots, N$, vanish. Hence, (C.36) simplifies to

$$\begin{aligned}\hat{r}_{LQ}(\mathbf{y}) &= \beta_0^* + \sum_{i,j} \beta_{i,j}^* y_i y_j^* \\ &= \beta_0^* + \boldsymbol{\beta}^H (\mathbf{y} \otimes \mathbf{y}^*). \end{aligned} \tag{C.37}$$

Accordingly, for the addressed ranging problem, the widely linear-quadratic MMSE estimator is obtained from (C.22) with $\boldsymbol{z} = \mathbf{y} \otimes \mathbf{y}^*$. In this particular case, inversion of a covariance matrix $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}$ of dimension $N^2 \times N^2$ instead of dimension $(2N + 3N^2) \times (2N + 3N^2)$ is needed.

Next, we compute the coefficients for the widely linear-quadratic MMSE estimator, which requires the moments of the entries of $\mathbf{y}$ up to fourth-order. Since $\mathbf{y}$ is circular, we only need to compute its even moments. From (C.27), (C.34), and (C.35), the second-order moments and cross moments are of the form

$$E[y_m y_n^*] = a_m a_n^* \mathcal{F}\{p(r)\} \left(\frac{f_m - f_n}{c} \right) [\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}]_{m,n} + \sigma^2 \delta_{m,n} \tag{C.38}$$

$$E[r y_m y_n^*] = a_m a_n^* \mathcal{F}\{r p(r)\} \left(\frac{f_m - f_n}{c} \right) [\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}]_{m,n} + \sigma^2 \delta_{m,n}. \tag{C.39}$$

The non-zero fourth-order moments of $\mathbf{y}$ have the form $E[y_i y_j^* y_m^* y_n]$. Inspecting (C.27) and dropping the cumulants that are zero, we obtain

$$E[y_i y_j^* y_m^* y_n] = \kappa(y_i y_j^* y_m^* y_n) + \kappa(y_i y_j^*)\kappa(y_n y_m^*) + \kappa(y_i y_m^*)\kappa(y_n y_m^*). \tag{C.40}$$

The second-order cumulants in (C.40) equal the second-order moments in (C.38). From (C.30) and (C.33), the remaining fourth-order cumulant reads

$$\kappa(y_i y_j^* y_m^* y_n) = a_i a_j^* a_m^* a_n \mathcal{F}\{p(r)\} \left(f_{\text{dif}}/c \right) 2 \int e^{-\jmath 2\pi f_{\text{dif}} \tau} \sigma_\alpha^4(\tau) \rho(\tau) \mathrm{d}\tau. \tag{C.41}$$

Thus, we have derived all moments necessary to calculate the coefficients of the widely linear-quadratic MMSE estimator in (C.37).

Inspection of expression (C.40) in combination with (C.23) shed some light on the properties of the estimator at extreme SNRs. When the SNR tends

Table C.1: Simulation Settings

OFDM system:
Bandwidth: 4.5 MHz, $N = 50$, $\Delta f = 15\,\text{kHz}$,
$T_\text{p} = T + T_\text{cp} = 66.7 + 5.4 = 72.1\,\mu\text{s}$, SNR $= \frac{E[\lvert a_i\rvert^2]}{\sigma^2}$,
Equal power and equal spacing pilot signal is used.
Channel parameters:
Homogenous Poisson point process: $\rho(\tau) = \rho_0$, $\lambda = 360\,\text{ns}$,
Average no. of paths: $\mu_L = \rho_0 T_\text{cp}$, $r \sim \mathcal{U}[0, 100]\text{m}$.
Results obtained from 10000 Monte Carlo trials are displayed. "Theo." is short for theoretical.

to infinity, i.e. the noise variance σ^2 approaches zero, (C.40) and (C.39) are constant. Using these results in (C.23), we can show that the RMSE approaches a constant when the SNR tends to infinity. At the other end of the scale, when the SNR tends to zero, $\boldsymbol{C_{zz}}$ is dominated by σ^2 and the RMSE approaches σ_r. In short,

$$\sqrt{E[|r - \hat{r}(\mathbf{y})|^2]} \to \begin{cases} \sigma_r & \text{for } \sigma^2 \to \infty \\ \text{constant} & \text{for } \sigma^2 \to 0. \end{cases}$$

5 Numerical Performance Evaluation

We first evaluate the performance of the approximate MAP, approximate MMSE and linear-quadratic MMSE estimators in a multi-path channel and contrast it with the performance of the non-coherent correlator-based estimator [15] [22] and the GAML estimator [14]. We omit the comparison with the multidimensional-search-based estimators [21] [18], because these estimators require access to L, which is assumed to be unknown.

5.1 Simulation Scenarios and Considered Estimators

Table C.1 summarizes the settings of the scenarios considered in the simulation. For simplicity, we adopt the homogeneous Poisson point process ($\rho(\tau) = \rho_0$) and assume that

$$\sigma_\alpha^2(\tau) = C \exp(-\frac{\tau}{\lambda}),$$

where C is a positive constant and λ denotes the root-mean-square (RMS) delay spread when the LOS component is removed from the channel response. With these assumptions, the delay power spectrum reads

$$P(\tau) = C \exp(-\frac{\tau}{\lambda})(\rho_0 + q^2\delta(\tau)). \tag{C.42}$$

Furthermore, we adopt a uniform prior for the range:

$$p(r) = \begin{cases} \frac{1}{r_{\max}} & \text{for } 0 < r \leq r_{\max} \\ 0 & \text{otherwise.} \end{cases}$$

This choice reflects the situation where the user terminal (to be localized) can appear at any distance to the base station within the interval $\mathcal{D} = [0, r_{\max}]$. Accordingly, the approximate MMSE estimator simplifies to

$$\hat{r}_{\text{AMMSE}}(\mathbf{y}) = \frac{1}{\int_0^{r_{\max}} e^{-\mathbf{y}^H \boldsymbol{C}_{\mathbf{y}}^{-1}(r)\mathbf{y}} \mathrm{d}r} \int_0^{r_{\max}} r e^{-\mathbf{y}^H \boldsymbol{C}_{\mathbf{y}}^{-1}(r)\mathbf{y}} \mathrm{d}r. \tag{C.43}$$

Based on the above assumptions on the multi-path channel, we now derive the analytical results that are needed to obtain the proposed estimators. We assume that $\int_{T_{\text{cp}}}^{-\infty} C e^{\frac{\tau}{\lambda}} \mathrm{d}\tau$ is negligible. Then, the entries of the covariance matrix $\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}$ in (C.8) read [14]

$$[\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}]_{ij} = q^2 C + \rho_0 g_{ij} \tag{C.44}$$

with

$$g_{ij} = C \frac{1 - e^{-(\jmath 2\pi(f_i - f_j) + \frac{1}{\lambda})T_{\text{cp}}}}{\jmath 2\pi(f_i - f_j) + \frac{1}{\lambda}} \quad i, j = 1, \ldots, N. \tag{C.45}$$

Note that $\mathbf{C}_{\tilde{\boldsymbol{\varepsilon}}}$ depends on the intensity function ρ_0 and the RMS delay spread λ. The values of ρ_0 and λ may be provided in the settings of an appropriate channel model for the propagation environment under consideration.

The fourth-order cumulants of $\mathbf{y}$ read

$$\kappa(y_i y_j^* y_m^* y_n) = a_i a_j^* a_m^* a_n \mathcal{F}\{p(r)\}\left(f_{\text{dif}}/c\right) 2\rho_0 C^2 \frac{1 - e^{(-\jmath 2\pi f_{\text{dif}} - \frac{2}{\lambda})T_{\text{cp}}}}{\jmath 2\pi f_{\text{dif}} + \frac{2}{\lambda}}. \tag{C.46}$$

The fourth-order moments of $\boldsymbol{\varphi}(r)$ in (C.46), see (C.34), read

$$\mathcal{F}\{p(r)\}\left(f_{\text{dif}}/c\right) = \begin{cases} \frac{1-e^{\varrho}}{\varrho} & i \neq j \\ 1 & i = j \end{cases} \tag{C.47}$$

with $\varrho = \jmath 2\pi f_{\text{dif}} r_{\max}/c$. Similarly, from (C.35) we have

$$\mathcal{F}\{rp(r)\}\left(f_{\text{dif}}/c\right) = \begin{cases} \frac{1}{\varrho}\left(\frac{1}{\varrho} + e^{\varrho}(r_{\max} + \frac{1}{\varrho})\right) & i \neq j \\ \mu_r & i = j \end{cases}. \tag{C.48}$$

Knowing the maximum range of r, (C.47) and (C.48) can be computed straightforwardly.

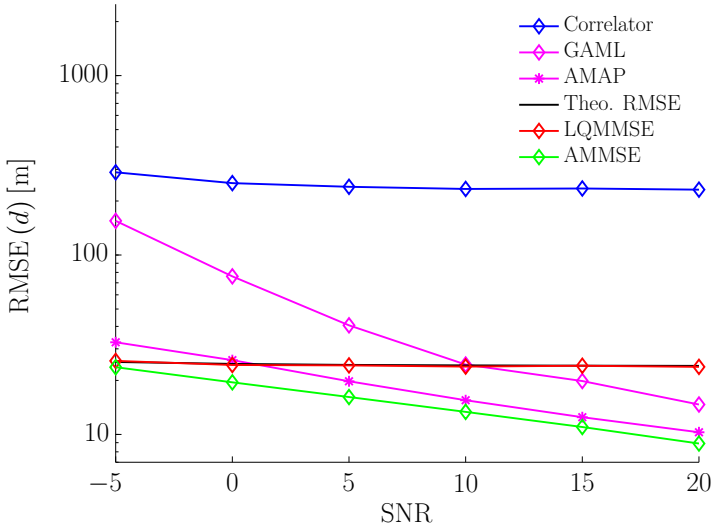


Fig. C.1: RMSEE versus SNR of the considered estimators: LOS condition and $\mu_L = 60$.

The above results are inserted in the corresponding estimators as needed. The argmax operation in the approximate MAP (AMAP) estimator (C.12) is carried out by first performing a coarse grid search to capture the main-lobe of the objective function followed by a refined search to improve the estimates. The integrals in the approximate MMSE (AMMSE) estimator (C.43) are computed numerically.

5.2 Ranging Accuracy Versus SNR in LOS Conditions

Fig. C.1 and Fig. C.2 report the simulated RMSE and bias of the investigated estimators. The GAML, approximate MAP, approximate MMSE, and widely linear-quadratic MMSE estimators significantly outperform the correlator-based estimator. Overall, the approximate MMSE estimator exhibits the best ranging performance in the considered SNR range. In the high SNR region, the approximate MAP and GAML estimators achieve higher ranging accuracy than the other considered estimators. They lose this advantage to the widely linear-quadratic MMSE estimator in the low SNR region. Because it exploits the prior information on r, the approximate MAP estimator outperforms the GAML estimator. In addition, the simulated RMSE of the widely linear-quadratic MMSE estimator coincides with the RMSE in (C.23). Fig. C.2 shows that the bias of all estimators decreases as the SNR increases. The widely linear-quadratic MMSE and approximate MMSE estimators are unbiased, while the other three exhibit positive biases.

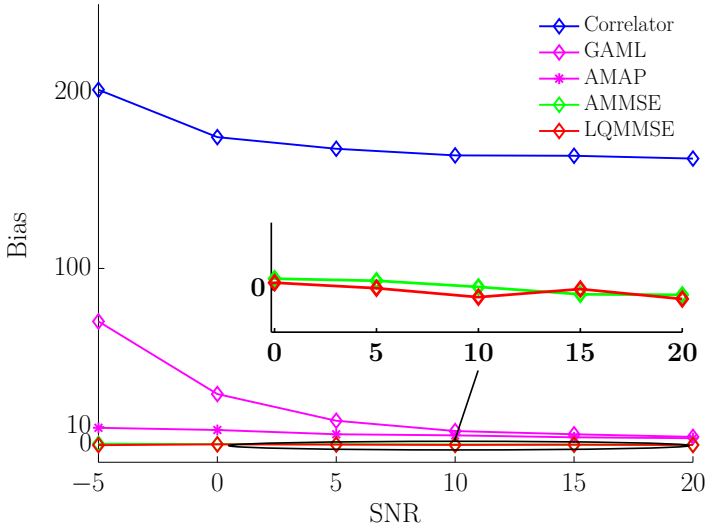


Fig. C.2: Bias versus SNR of the considered estimators: LOS condition and $\mu_L = 60$.

5.3 Performance Evaluation Under Different Channel Settings

Fig. C.3 reports the simulated RMSEs versus the average number of path components $\mu_L = \rho_0 T_{\text{cp}}$ for the LOS condition. We observe again that the GAML, approximate MAP, approximate MMSE and widely linear-quadratic MMSE estimators significantly outperform the correlator-based estimator, with the approximate MMSE estimator achieving the highest ranging accuracy. As the average number of paths increases, the performance of the approximate MAP, approximate MMSE, and widely linear-quadratic MMSE estimators deteriorate, while the performance of the GAML estimator improves. The simulated RMSE of the widely linear-quadratic MMSE estimator coincides with the RMSE in (C.23).

Fig. C.4 reports the simulated RMSEs in the NLOS condition. Contrary to what is observed in the LOS condition, the RMSE decreases as μ_L increases. Compared to the correlator-based and GAML estimators, the widely linear-quadratic MMSE estimator achieves a promising performance gain, in particular when the average number of paths is small. When the average number of paths is large, the approximate MMSE estimator shows the best ranging accuracy among the considered estimators. This is a consequence of the fact that the Gaussian approximation of the distribution of $\mathbf{y}$ becomes more accurate. When the average number of paths is small, leading to a more pronounced deviation of the distribution of $\mathbf{y}$ from the Gaussian approximation, the fourth-order mo-

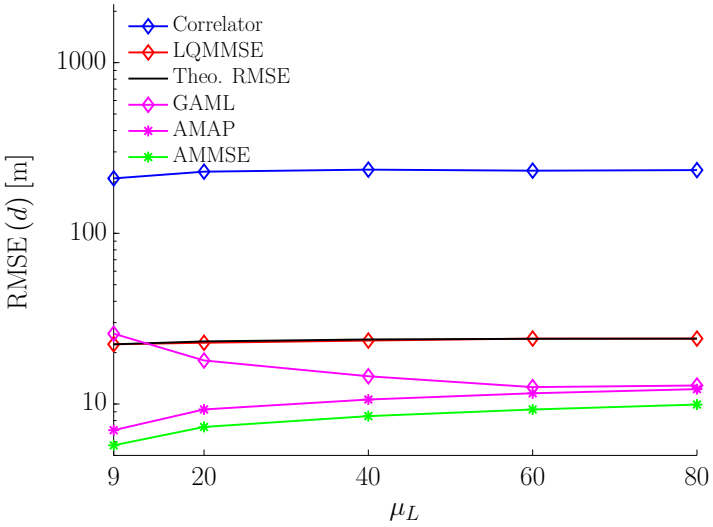


Fig. C.3: RMSEE versus average number of paths μ_L of the considered estimators: LOS condition and SNR = 20 dB.

ments have a larger impact on the ranging accuracy. As a result, the widely linear-quadratic MMSE estimator outperforms the approximate MMSE estimator. Finally, because it utilizes the prior knowledge on r, the approximate MAP estimator achieves higher ranging accuracy than the GAML estimator.

5.4 Performance Comparison in Model Mismatch

The performance of the proposed estimators is anticipated to degrade if there is a mismatch between the settings of the parameters of the channel in which the estimators effectively operate and the selected channel parameter settings used in these estimators. From (D.23), the intensity function ρ_0, the RMS delay spread λ, and the LOS indicator q are the parameters governing the delay power spectrum of the channel. In Section 5.3, we have investigated the impact of the average number of path components μ_L, and therefore of ρ_0, on the performance of the proposed estimators. Here, we report the impact of having mismatches on the RMS delay spread λ and the LOS indicator q. For the RMS delay spread, we say that there is a mismatch if the selected setting of λ used to derive the considered estimators is different from that setting of the channel in which the estimators effectively operate. This reflects the situation that either the estimate of λ or the information on λ that we obtained from the settings of the channel model has some uncertainties. For the LOS indicator, we say that there is a mismatch if a LOS (NLOS) condition prevails

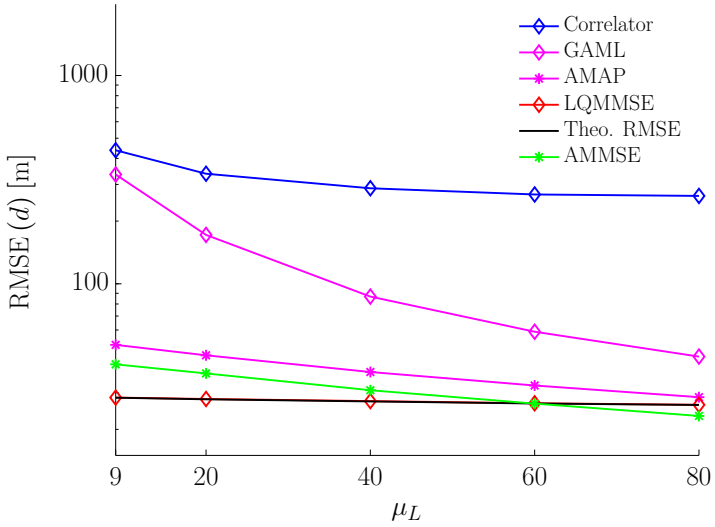


Fig. C.4: RMSEE versus the average number of paths μ_L of the considered estimators: NLOS condition and SNR = 20 dB.

in the channel, while the value of q used in the estimators correspond to a NLOS (LOS) condition. Otherwise there is a match. This study has practical implications: since detecting LOS and NLOS conditions can be difficult in many scenarios, it allows for testing the robustness of the proposed estimators against such channel state mismatches.

We show the impact of a mismatch on λ to the RMSE of the estimators in Fig. C.7 (LOS condition) and Fig. C.8 (NLOS condition). In the LOS condition, the mismatch causes performance loss. However, this loss is negligible when the widely linear-quadratic MMSE estimator is employed. The approximate MAP and approximate MMSE estimators show a noticeable loss, but they still outperform the widely linear-quadratic MMSE estimator. Moreover, the smaller the RMS delay spread, the smaller the loss, since a smaller RMS delay spread leads to fewer path components with large delays appearing in the channel response. In the NLOS condition, the widely linear-quadratic MMSE estimator is still quite robust against the mismatch. Despite the mismatch on λ, the ranging accuracy increases when the RMS delay spread decreases. This implies that even if the value of λ used in the estimators is wrong, the approximate MAP and MMSE estimators still lead to a lower RMSE due to the better channel condition.

The impact of the LOS/NLOS mismatch is shown in Fig. C.5 and Fig. C.6. It appears that both approximate MAP and widely linear-quadratic MMSE

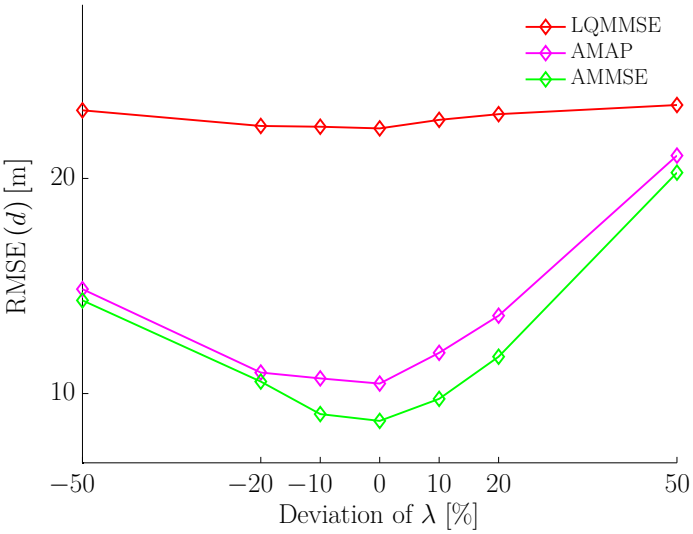


Fig. C.5: RMSEE versus the deviation of λ that is used in the channel as compared to that is assumed in the estimators: LOS condition, SNR = 20 dB, and $\mu_L = 60$.

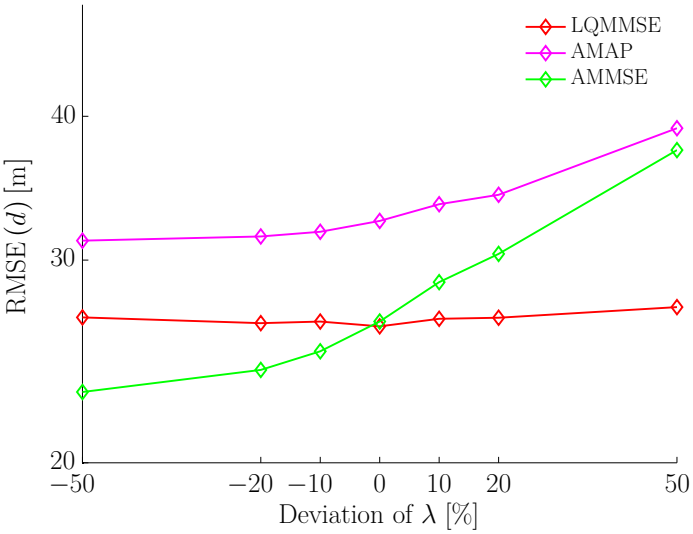


Fig. C.6: RMSEE versus the deviation of λ that is used in the channel as compared to that is assumed in the estimators: NLOS condition, SNR = 20 dB, and $\mu_L = 60$.

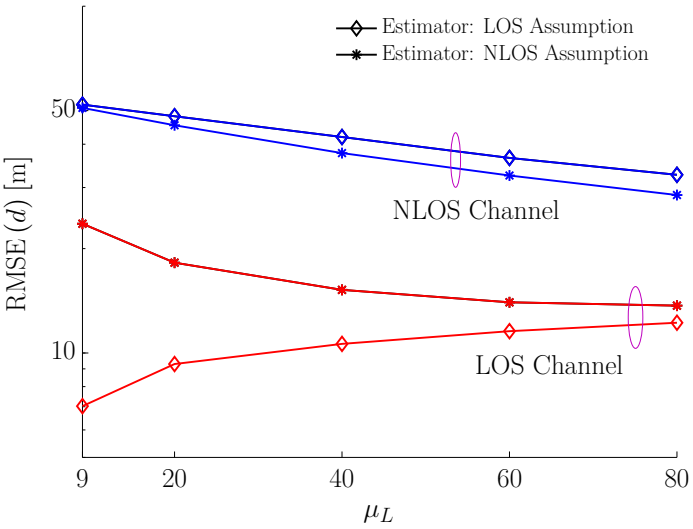


Fig. C.7: RMSEE of the approximate MAP estimator versus average number of paths μ_L for SNR = 20 dB. Blue and red curves indicate the real channel conditions. "Estimator: LOS Assumption" shows the RMMSE of estimator (C.12) assuming that the real propagation channel is in a LOS condition while "Estimator: NLOS Assumption" denotes the RMMSE of estimator (C.12) assuming that the real channel is in a NLOS condition.

estimators benefit from the LOS condition, which leads to a lower RMSE. In case of a mismatch, both estimators perform worse than when there is a match. This is particularly noticeable in the LOS condition with small μ_L. Except for the matched case when a LOS condition prevails, the RMSE decreases as μ_L increases. We remark that the approximate MMSE and approximate MAP estimators behave similarly. Therefore, we omit to include their performance plots.

6 Conclusion

The proposed direct ranging methods bypass the first-path detection problem and obviate the requirements of knowing the number of path components in the channel response and of any separability condition of these components. These methods rely on a channel model formulated via a point process approach, which allows for computing the required moments of the channel response. If the first- and second-order moments of the received signal are available, the approximate MAP and MMSE estimators can be efficiently evaluated. If we further have access to up to $2p$th-order moments, the unbiased pth-order polynomial MMSE estimator can be applied. By using Turin's classical channel

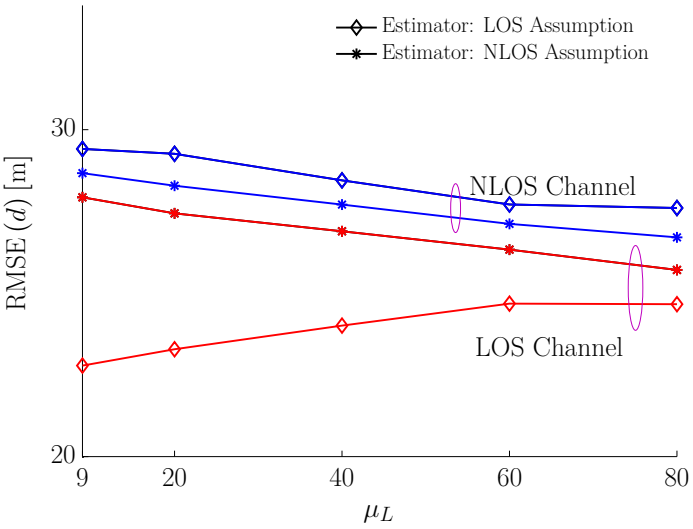


Fig. C.8: RMSEE of the widely linear-quadratic MMSE estimator versus average number of paths μ_L for SNR = 20 dB. Blue and red curves indicate the real channel conditions. "Estimator: LOS Assumption" shows the RMMSE of estimator (C.22) assuming that the real propagation channel is in a LOS condition while "Estimator: NLOS Assumption" denotes the RMMSE of estimator (C.22) assuming that the real channel is in a NLOS condition.

model, all moments of the received signal can be computed analytically via Campbell's theorem.

Remarkably, the standard linear MMSE estimator ($p = 1$) is inapplicable in this case, whereas the widely linear-quadratic MMSE estimator ($p = 2$) can be employed. Our findings show that if a prior information on the range is available, Bayesian estimators, such as the approximate MAP, the approximate MMSE, and the widely linear-quadratic MMSE estimators significantly outperform the non-coherent correlator-based estimator. Furthermore, simulation results demonstrate that the proposed estimators are fairly robust against model mismatches.

For the pth-order polynomial MMSE estimator, the required moments of the observation vector are needed, not its pdf. In addition, the RMSE of the polynomial MMSE estimator can be analytically computed. This is not the case for the approximate MAP and MMSE estimators. The performance of these two estimators are inherently limited by the underlying Gaussian approximation. When a multi-path channel with low average number of path components is encountered, the Gaussian approximation of the distribution of $\mathbf{y}$ is poor. In this particular case or in similar cases, a straightforward way to improve the performance of the two estimators is to investigate more accurate approximations that take into account higher-order moments of the received signal.

An aspect not considered in this contribution that can be the subject of future investigations is the distribution of range errors. This distribution is widely used to derive range-based localization algorithms. As with other ranging methods, the exact distribution of the range errors of the proposed estimators is not known. Thus, one must resort to empirical models of the pdf of the range errors. Since direct ranging methods do not rely on first-path detection, potential range errors resulting from missed first-path detection, which are often considered in error models, are avoided. This rationale leads us to conjecture that direct ranging techniques give rise to differently distributed range errors compared to existing ranging methods.

A Inversion of $\boldsymbol{C_{zz}}$

For circular $\mathbf{y}$, the covariance matrix in (C.18) with $p = 2$ is of the form

$$\boldsymbol{C_{zz}} = \begin{bmatrix} \mathbf{C_y} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C_y^T} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{C_{z_1}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{C_{z_2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{C_{z_3}} \end{bmatrix},$$

where

$$\boldsymbol{C}_{\boldsymbol{z}_i} = E[(\boldsymbol{z}_i - E[\boldsymbol{z}_i])(\boldsymbol{z}_i - E[\boldsymbol{z}_i])^H]$$

with

$$\boldsymbol{z}_1 = \mathbf{y} \otimes \mathbf{y}, \boldsymbol{z}_2 = \mathbf{y}^* \otimes \mathbf{y}^*, \boldsymbol{z}_3 = \mathbf{y} \otimes \mathbf{y}^*. \tag{C.49}$$

Obviously, $\boldsymbol{z}_1 = \boldsymbol{z}_2^*$ leading to $\boldsymbol{C}_{\boldsymbol{z}_2} = \boldsymbol{C}_{\boldsymbol{z}_1}^T$. The block diagonal structure of $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}$ greatly reduces the computational complexity of inverting this matrix. The complexity of computing $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1}$ amounts to that of inverting the covariance matrix $\mathbf{C}_\mathbf{y}$, which contains the second-order moments, and the matrices $\boldsymbol{C}_{\boldsymbol{z}_1}$ and $\boldsymbol{C}_{\boldsymbol{z}_3}$, which contain the fourth-order moments of $\mathbf{y}$. Notice that $\boldsymbol{C}_{\boldsymbol{z}_1}$ and $\boldsymbol{C}_{\boldsymbol{z}_3}$ have the same entries but ordered differently.

B Widely Linear MMSE and Linear MMSE Estimator

The widely linear MMSE estimator for a complex signal $\mathbf{y}$ reads [17]

$$\hat{r}(\mathbf{y}) = \beta_0^* + \mathbf{h}^H\mathbf{y} + \mathbf{g}^H\mathbf{y}^*. \tag{C.50}$$

Thus $\boldsymbol{\beta}^H = [\mathbf{h}^H\ \mathbf{g}^H]$ and $\boldsymbol{z} = [\mathbf{y}^T\mathbf{y}^H]^T$. Referring to (C.21), $\boldsymbol{C}_{r\boldsymbol{z}}$ needs to be computed in order to obtain $\boldsymbol{\beta}$. Since $E[\boldsymbol{z}] = \mathbf{0}$, the cross-covariance matrix $\boldsymbol{C}_{r\boldsymbol{z}}$ equals $E[r\boldsymbol{z}^H]$. Inserting (B.1), the entry $E[ry_i^{\diamondsuit_i}]$ of $E[r\boldsymbol{z}^H]$ reads

$$E[ry_i^{\diamondsuit_i}] = E[r\varphi_i^{\diamondsuit_i}(r)]\underbrace{E[\tilde{\varepsilon}_i^{\diamondsuit_i}]}_{=0} + \underbrace{E[n_i^{\diamondsuit_i}]}_{=0} = 0, \tag{C.51}$$

which follows from the assumption that $\tilde{\varepsilon}_i$ and r are independent. Thus, $\boldsymbol{\beta}^H = \mathbf{0}$ leading to $\boldsymbol{\beta}_0 = \mu_r$. The widely linear MMSE estimator

$$\hat{r}(\mathbf{y}) = \mu_r$$

is therefore independent of the observed data. It only depends on the prior knowledge of r. Therefore, this estimator is inapplicable to our problem.

Setting either $\mathbf{h}$ or $\mathbf{g}$ to be zero, the estimator is commonly referred to as the linear MMSE estimator [3]. Following the same procedure as to obtain the widely linear MMSE estimator, we show that the linear and widely linear MMSE estimators coincide as a consequence of $\mathbf{y}$ being circular [17]. Therefore, we do not distinguish between the widely linear MMSE estimator and the linear MMSE estimator.

C Linear Quadratic MMSE Estimator

The general form of a widely linear-quadratic MMSE estimator for a complex signal model reads

$$\begin{aligned}\hat{r}_{LQ}(\mathbf{y}) &= \beta_0^* + \sum_i g_i^* y_i + \sum_i h_i^* y_i^* \\ &\quad + \sum_{i,j} \gamma_{i,j}^* y_i y_j + \sum_{i,j} \vartheta_{i,j}^* y_i^* y_j^* + \sum_{i,j} \beta_{i,j}^* y_i y_j^*. \end{aligned} \tag{C.52}$$

It follows from (C.51) and the circularity of $\mathbf{y}$ that $g_i = h_i = 0$, $i = 1, \ldots, N$. Thus, (C.52) simplifies to

$$\hat{r}_{LQ}(\mathbf{y}) = \beta_0^* + \sum_{i,j} \gamma_{i,j}^* y_i y_j + \sum_{i,j} \vartheta_{i,j}^* y_i^* y_j^* + \sum_{i,j} \beta_{i,j}^* y_i y_j^*.$$

To proceed, we formulate $\boldsymbol{z} = [\boldsymbol{z}_1^T \ \boldsymbol{z}_2^T \ \boldsymbol{z}_3^T]^T$ with $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, and $\boldsymbol{z}_3$ given in (C.49). We define the column vector of the corresponding filter coefficients $\boldsymbol{\beta} = [\boldsymbol{\gamma}^T \ \boldsymbol{\vartheta}^T \ \boldsymbol{\beta}_1^T]^T$.

First, we compute the coefficients associated to $\boldsymbol{z}_2$. Applying the orthogonality principle (C.20), we obtain

$$E[r\boldsymbol{z}_1^H] = \beta_0^* E[\boldsymbol{z}_1^H] + \boldsymbol{\gamma}^H E[\boldsymbol{z}_1 \boldsymbol{z}_1^T] + \boldsymbol{\vartheta}^H E[\boldsymbol{z}_1 \boldsymbol{z}_2^T] + \boldsymbol{\beta}_1^H E[\boldsymbol{z}_1 \boldsymbol{z}_3^T]. \tag{C.53}$$

We can show that the entries of $E[r\boldsymbol{z}_1^H]$ vanish:

$$E[r y_i y_j] = E[r\varphi_i(r)\varphi_j(r)] \underbrace{E[\tilde{\varepsilon}_i \tilde{\varepsilon}_j]}_{=0} + \underbrace{E[n_i n_j]}_{0} = 0.$$

This results from $\tilde{\boldsymbol{\varepsilon}}$ being circular and $\mathbf{n}$ being white. Since the vector $\mathbf{y}$ is circular, $E[\boldsymbol{z}_1^H] = \mathbf{0}$. Furthermore, $E[\boldsymbol{z}_1 \boldsymbol{z}_1^T] = E[\boldsymbol{z}_1 \boldsymbol{z}_3^T] = \mathbf{0}$. Consequently, (C.53) simplifies to

$$\boldsymbol{\vartheta}^H E[\boldsymbol{z}_1 \boldsymbol{z}_2] = \boldsymbol{\vartheta}^H E[\boldsymbol{z}_1 \boldsymbol{z}_1^H] = \mathbf{0}. \tag{C.54}$$

Since the matrix $E[\boldsymbol{z}_1 \boldsymbol{z}_1^H]$ is positive definite, by (C.54) $\boldsymbol{\vartheta} = \mathbf{0}$. Following the same procedure, we obtain $\boldsymbol{\gamma} = \mathbf{0}$. Consequently, the widely linear-quadratic MMSE estimator is of the form

$$\hat{r}_{LQ}(\mathbf{y}) = \beta_0^* + \sum_{i,j} \beta_{i,j}^* y_i y_j^*. \tag{C.55}$$

Acknowledgment

This work was supported in part by the EU FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306).

References

[1] A. Hero. Timing estimation for a filtered Poisson process in Gaussian noise. *IEEE Trans. Inf. Theory*, 37(1):92–106, Jan. 1991.

[2] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen. Channel parameter estimation in mobile radio environments using the SAGE algorithm. *IEEE J. Sel. Areas Commun.*, 17(3):434–450, Mar. 1999.

[3] B. Picinbono and P. Chevalier. Widely linear estimation with complex data. *IEEE Trans. Signal Process.*, 43(8):2030–2033, Aug. 1995.

[4] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win. Ranging with ultrawide bandwidth signals in multi-path environments. *Proc. IEEE*, 97(2):404–426, 2009.

[5] G. L. Turin, F. D. Clapp, T. L. Johnston, S. B. Fine, and D. Lavry. A statistical model of urban multipath propagation. *IEEE Trans. Veh. Technol.*, 21(1):1–9, 1972.

[6] G. Steinbock, T. Pedersen, B. H. Fleury, W. Wang and R. Raulefs. Distance dependent model for the delay power spectrum of in-room radio channels. *IEEE Trans. Antennas Propag.*, 61(8):4327–4340, Aug. 2013.

[7] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory - Part III.* John Wiley & Sons, 2001.

[8] J. Eriksson, E. Ollila, and V. Koivunen. Essential statistics and tools for complex random variables. *IEEE Trans. Signal Process.*, 58(10):5400–5408, Oct. 2010.

[9] J. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proc. IEEE*, 79(3):278–305, Mar. 1991.

[10] J. N. Kapur. *Maximum-entropy Models in Science and Engineering.* Wiley, 1989.

[11] M. L. Jakobsen. *Modeling and Analysis of Stochastic Radio Channels.* PhD thesis, Aalborg University, 2014.

[12] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice-Hall PTR, 1998.

[13] J. Kingman. *Poisson Processes.* Oxford studies in probability. Clarendon Press, 1992.

[14] L. Jing, T. Pedersen, and B. Fleury. Direct ranging in multi-path channels using OFDM pilot signals. In *15th IEEE Int. Symposium on Signal Process. Advances in Wireless Commun.*, pages 150–154, June 2014.

[15] C. Mensing. *Location Determination in OFDM Based Mobile Radio Systems*. PhD thesis, Technische Universität München, 2013.

[16] P. Bondon, M. Benidir, and B. Picinbono. A nonlinear approach to estimate the amplitude of a signal. In *IEEE International Conf. on Acoustics, Speech, and Signal Process.*, volume 5, pages 301–304, Mar. 1992.

[17] P. Schreier and L. Scharf. *Statistical Signal Processing of Complex-Valued Data*. Cambridge University Press, 2010.

[18] R. Adam and P. A. Hoeher. Semi-blind channel estimation for joint communication and positioning. *10th Workshop on Positioning Navigation and Commun.*, pages 1–5, 2013.

[19] S. Marano, W. Gifford, H. Wymeersch, and M. Win. NLOS identification and mitigation for localization based on UWB experimental data. *IEEE J. Sel. Areas Commun.*, 28(7):1026–1035, 2010.

[20] A. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location. *IEEE Signal Process. Mag.*, 22(4):24–40, 2005.

[21] T. Wang, Y. Shen, S. Mazuelas, H. Shin and M. Z. Win. On OFDM ranging accuracy in multipath channels. *IEEE Systems Journal*, PP(99):1–11, 2013.

[22] D. Wang, M. Fattouche, and F. Ghannouchi. Fundamental limit of OFDM range estimation in a separable multipath environment. *Circuits, Systems, and Signal Processing*, 31(3), 2012.

Paper D

Bayesian Ranging for Radio Localization with and without Line-of-Sight Detection

Lishuai Jing, Troels Pedersen, Bernard H. Fleury

The paper has been published in the
IEEE International Conference on Communications (ICC) Workshop on Advances in Network Localization and Navigation (ANLN)
Jun. 2015.

The layout has been revised.

Abstract

We consider Bayesian ranging methods for localization in wireless communication systems. Based on a channel model and given priors for the range and the line-of-sight (LOS) condition, we propose range estimators with and without LOS detection. Since the pdf of the received frequency-domain signals is unknown, we approximate the maximum-a-posteriori (MAP) and the minimum mean-squared error (MMSE) estimators. The promising ranging accuracy obtained with the proposed estimators is demonstrated by Monte Carlo simulations. We observe that the approximate MMSE estimators outperform the approximate MAP estimators. In addition, we find that including LOS detection in the approximate estimators, while adding a higher computational complexity, has no major impact on the ranging performance.

1 Introduction

Having accurate localization capability is increasingly important for wireless communication systems [14] [6]. One approach to increase localization performance is to rely on high precision ranging techniques [3]. State-of-the-art ranging techniques based on, for example, the received signal strength, angle-of-arrival, time-of-arrival, time-difference-of-arrival, etc, may be sensitive to line-of-sight (LOS) conditions [6] [16]. Therefore, accounting for the unknown LOS or non-LOS (NLOS) conditions is an issue considered in many ranging and localization techniques [7] [9].

To tackle this issue, one approach is to rely on LOS identification techniques. Such a approach is reliable provided that the signal bandwidth and signal-to-noise ratio (SNR) are sufficiently large [16] [22] [10] [5] [17] [2]. Existing LOS identification techniques include methods based on machine-learning [17] [4] and hypothesis-testing [7] [5] [6]. The LOS identification step labels range estimates as "LOS" or "NLOS" to facilitate the localization algorithms [21] [4]. The rational is that if the LOS condition can be correctly identified, this information can be used to improve the ranging and localization accuracy. However, in communication systems with limited bandwidth and SNR, the LOS detector may be unreliable [16].

Instead of identifying and mitigating NLOS range estimates, direct ranging, which infers the range parameter directly from the received signal, can be potentially applied. Direct ranging methods have been proposed in [12, 13] for bypassing a related problem, i.e. the first-path detection. These methods rely on a channel model formulated via a point process to compute the required moments of the received signal. In [13], an approximate maximum-likelihood ranging method using the first- and second-order moments of the received signal has been presented. Using the prior distribution of the range,

Bayesian estimators, including approximate maximum a posteriori (MAP) and minimum-mean-square-error (MMSE) estimators and a pth-order MMSE polynomial estimator, are proposed in [12]. In contrast to the methods in [20] [15], direct ranging operates without knowledge of the number of multi-path components in the channel response and separability condition on these components. Although the methods in [12, 13] still rely on LOS state information, the principle of ranging without estimating intermediate parameters seems promising.

In the present contribution, we propose Bayesian ranging methods with and without LOS detection for multi-path channels. Inspired by the direct ranging principle, we make use of a channel model to approximate pdfs of the received signal. In addition, we incorporate prior information on the range and the LOS condition. For this setup, we propose and evaluate approximate MAP and MMSE estimators. In addition, we derive variants of these estimators with approximate MAP and Bayes decision rules for LOS detection. We test the performance of the proposed methods by means of Monte Carlo simulations of an OFDM system with limited bandwidth.

2 Signal and Channel Model

We address the problem of estimating the range parameter r directly from the received signal vector $\mathbf{y} = [y_1, \ldots, y_N]^T$ obtained at frequencies $f_1, \ldots, f_N$. We follow the Bayesian approach and consider the range r to be a random variable with a priori pdf $p(r)$. Assuming the channel to be time-invariant with additive noise, we write

$$\mathbf{y} = \mathbf{A}\mathbf{h}(r) + \mathbf{n}, \tag{D.1}$$

where $\mathbf{A} = \mathrm{diag}\{a_1, \ldots, a_N\}$ is a diagonal matrix containing the known pilot symbol, $\mathbf{h}(r)$ denotes the range-dependent frequency-domain channel response, and $\mathbf{n}$ is a white circular complex Gaussian noise vector with component variance σ^2.

As in [13] and [12], we decompose $\mathbf{h}(r)$ as the Hadamard product of a range-dependent factor $\boldsymbol{\varphi}(r)$ and a range-independent factor $\boldsymbol{\xi}$:

$$\mathbf{h}(r) = \boldsymbol{\varphi}(r) \odot \boldsymbol{\xi} \tag{D.2}$$

with

$$\boldsymbol{\varphi}(r) = [\varphi_1, \ldots, \varphi_N]^T, \quad \varphi_n = e^{-\jmath 2\pi f_n \frac{r}{c}},$$

where $\jmath = \sqrt{-1}$ and c is the speed of light. Unlike [13] and [12], we here consider the case of a multi-path channel in which LOS propagation occurs with probability p_{LOS}. Thus we write $\boldsymbol{\xi}$ as a superposition of a LOS term and a multi-path term

$$\boldsymbol{\xi} = q\alpha_0\mathbf{1} + \boldsymbol{\varepsilon}, \tag{D.3}$$

where $\mathbf{1}$ denotes an all-ones vector. The random LOS indicator q takes value one with probability p_{LOS} and zero otherwise. The complex gain of the LOS term is denoted by α_0. The multi-path term $\boldsymbol{\varepsilon} = [\varepsilon_1, \ldots, \varepsilon_N]^T$ has entries

$$\varepsilon_n = \sum_{l=1}^{L} \alpha_l e^{-j2\pi f_n \tau_l}, \quad n = 1, \ldots, N, \tag{D.4}$$

where α_l is the complex gain and τ_l is the excess delay of path l with respect to the LOS delay $\frac{r}{c}$. The random excess delays form a point process $\mathcal{T} = \{\tau_1, \tau_2, \ldots\}$ with intensity function $\rho(\tau)$ whose shape controls the average number of points in $\mathcal{T}$ per time unit. By convention, we set the delay associated to the LOS component to be zero, i.e. $\tau_0 = 0$. Depending on the specific point process assumed, the number $L = |\mathcal{T}|$ of multi-path components may be random and potentially infinite. We further assume that

$$E[\alpha_l|\tau_l] = 0, \;\; E[\alpha_l \alpha_{l'}^*|\tau_l, \tau_{l'}] = \begin{cases} \sigma_\alpha^2(\tau_l), & l = l' \\ 0, & \text{otherwise,} \end{cases} \tag{D.5}$$

where $\sigma_\alpha^2(\tau_l)$ denotes the expected power of a path component with delay τ_l. With these definitions, the delay power spectrum of the considered channel model is of the form

$$P(\tau) = E[P(\tau|q)|q] \tag{D.6}$$

where $P(\tau|q)$ is the conditional delay power spectrum [13]

$$P(\tau|q) = \sigma_\alpha^2(\tau)(\rho(\tau) + q^2\delta(\tau)) \tag{D.7}$$

with δ denoting the Dirac delta function. Thus, $P(\tau) = \sigma_\alpha^2(\tau)(\rho(\tau) + p_{\mathrm{LOS}}\delta(\tau))$.

3 Estimation of Range

3.1 Approximate Likelihood Function

Standard Bayesian estimators such as MAP and MMSE estimators necessitate the computation of the posterior pdf $p(r|\mathbf{y})$. For a specific estimation problem, this pdf may be known directly or alternatively computed via Bayes Theorem, provided that the likelihood function $p(\mathbf{y}|r)$ is known. For the problem described in Section 2, it is most convenient to work with the likelihood function, which can be expressed as

$$p(\mathbf{y}|r) = \sum_{q=0}^{1} p(\mathbf{y}|r, q)p(q), \tag{D.8}$$

where $p(q)$ denotes the probability mass function of q. Unfortunately, for the case considered, the two likelihood functions $p(\mathbf{y}|r, q)$ and $p(\mathbf{y}|r)$ are unknown

and therefore we resort to approximations. Here, we consider two different approximations for $p(\mathbf{y}|r)$.

To derive the first approximation, we follow the same approach as in [13] and [12]: we approximate the likelihood function $p(\mathbf{y}|r,q)$ as a Gaussian pdf $p_{\mathrm{G}}(\mathbf{y}|r,q)$ with the same first and second moments, i.e. with mean zero and covariance

$$\mathbf{C}_{\mathbf{y}|r,q} = E[\mathbf{y}\mathbf{y}^H|r,q] = \mathbf{A}\boldsymbol{\Phi}(r)\mathbf{C}_{\boldsymbol{\xi}|q}\boldsymbol{\Phi}^H(r)\mathbf{A}^H + \sigma^2\mathbb{I}, \tag{D.9}$$

where $\boldsymbol{\Phi}(r) = \mathrm{diag}\{\boldsymbol{\varphi}(r)\}$, $\mathbb{I}$ denotes the identity matrix, and $\mathbf{C}_{\boldsymbol{\xi}|q} = E[\boldsymbol{\xi}\boldsymbol{\xi}^H|q]$ with the (m,n)th entry computed as

$$[\mathbf{C}_{\boldsymbol{\xi}|q}]_{mn} = \mathcal{F}\{P(\tau|q)\}(f_m - f_n). \tag{D.10}$$

Here, $\mathcal{F}$ denotes the Fourier transform. Inserting $p_{\mathrm{G}}(\mathbf{y}|r,q)$ for $p(\mathbf{y}|r,q)$ in (D.8), we obtain the Gaussian mixture

$$p_{\mathrm{GM}}(\mathbf{y}|r) = \sum_{q=0}^{1} p_{\mathrm{G}}(\mathbf{y}|r,q)p(q). \tag{D.11}$$

In the second approximation, we replace $p(\mathbf{y}|r)$ directly by a Gaussian $p_{\mathrm{G}}(\mathbf{y}|r)$ with the same first and second moments as $\mathbf{y}|r$, i.e. with mean zero and covariance

$$\mathbf{C}_{\mathbf{y}|r} = E[\mathbf{C}_{\mathbf{y}|r,q}] \tag{D.12}$$

in which $\mathbf{C}_{\boldsymbol{\xi}} = E[\mathbf{C}_{\boldsymbol{\xi}|q}]$ can be straightforwardly computed.

Evaluation of $p_{\mathrm{G}}(\mathbf{y}|r)$ and $p_{\mathrm{GM}}(\mathbf{y}|r)$ requires calculation of determinants and inverses of the matrices defined in (D.9) and (D.12). Following the same line of arguments as in [13], these computation tasks simplify since the determinants do not depend on r and inversion of the involved matrices can be carried out efficiently.

The accuracy of the above approximations depends on the specific parameter settings of the channel model. As an example, the Gaussian approximation may be inaccurate if the average number of path components in the multipath channel, see (D.4), is small or the delay power spectrum exhibits a fast exponential decay. In the other extreme where the delay power spectrum is a constant and the average number of path components is high, the Gaussian approximation is well justified. Consequently, the accuracy of the estimators derived from the proposed approximations should be assessed, e.g. via Monte Carlo simulations.

3.2 Approximate MAP Ranging

The MAP estimator for r, defined as

$$\hat{r}_{\mathrm{MAP}}(\mathbf{y}) = \arg\max_r p(\mathbf{y}|r)p(r), \tag{D.13}$$

cannot be computed since $p(\mathbf{y}|r)$ is unknown. Therefore, we propose to approximate it by replacing $p(\mathbf{y}|r)$ with either $p_{GM}(\mathbf{y}|r)$ or $p_{\mathrm{G}}(\mathbf{y}|r)$ defined above. Accordingly, we define two approximate MAP estimators:

$$\hat{r}_{\mathrm{AMAP,GM}}(\mathbf{y}) = \arg\max_{r} p_{\mathrm{GM}}(\mathbf{y}|r)p(r), \tag{D.14}$$

$$\hat{r}_{\mathrm{AMAP,G}}(\mathbf{y}) = \arg\max_{r} p_{\mathrm{G}}(\mathbf{y}|r)p(r). \tag{D.15}$$

In (D.14) and (D.15), we marginalized over q and therefore LOS detection is not needed. Alternatively, we can obtain the range by detecting the LOS condition first. This results in an approximate MAP estimator for r:

$$\hat{r}_{\mathrm{AMAP,Dec}}(\mathbf{y}) = \arg\max_{r} p_{\mathrm{G}}(\mathbf{y}|r, \hat{q})p(r), \tag{D.16}$$

with $\hat{q}$ denoting the approximate MAP decision rule

$$\hat{q}(\mathbf{y}) = \arg\max_{q} p(q) \int p_{\mathrm{G}}(\mathbf{y}|r, q)p(r)\mathrm{d}r. \tag{D.17}$$

Computation of (D.16) and (D.17) is a two-step procedure with a LOS detection step followed by a ranging step. However, it is unclear if this additional complexity due to the LOS detector translates into improved ranging accuracy since the involved Gaussian approximations may undermine the performance of (D.16) and (D.17). In Section 4, we carry out a simulation study to answer this question.

Depending on the choice of prior and delay power spectrum, the optimization in (D.14)–(D.17) may require numerical procedures. We remark that to numerically evaluate the objective functions, it is necessary to invert the corresponding covariances defined in (D.9) and (D.12) for each value of r. As already shown (see [13]), this inversion can be simplified using eigenvalue decomposition.

3.3 Approximate MMSE Ranging

For the ranging problem, the MMSE estimator is given by

$$\hat{r}_{\mathrm{MMSE}}(\mathbf{y}) = \arg\min_{r'} E[(r - r')^2|\mathbf{y}] = E[r|\mathbf{y}], \tag{D.18}$$

where the expectation is taken over the unknown pdf $p(r|\mathbf{y})$.

Using the approximations for $p(r|\mathbf{y})$ in Section 3.1, we obtain approximate MMSE estimators:

$$\hat{r}_{\mathrm{AMMSE,GM}}(\mathbf{y}) = E_{p_{\mathrm{GM}}}[r|\mathbf{y}], \tag{D.19}$$

$$\hat{r}_{\mathrm{AMMSE,G}}(\mathbf{y}) = E_{p_{\mathrm{G}}}[r|\mathbf{y}], \tag{D.20}$$

where the expectations are taken over $p_{\mathrm{GM}}(r|\mathbf{y})$ and $p_{\mathrm{G}}(r|\mathbf{y})$ respectively.

The performance of the estimators (D.19) and (D.20) is essentially limited by the involved approximations. These estimators are therefore not optimal in a particular sense. Better performing estimators could potentially be obtained by invoking more accurate approximations. One candidate improvement provided a reliable detection of the LOS condition is to use separate approximations for the LOS and NLOS cases. Here, we consider Bayes' decision rule in combination with the approximate MMSE estimator defined in (D.20) where $p_{\mathrm{LOS}} = 1$ when LOS is detected and zero otherwise:

$$\hat{r}_{\mathrm{AMMSE,D}} = E_{\mathrm{pG}}[r|\mathbf{y}, \hat{q}]. \tag{D.21}$$

Bayes' decision rule for q reads

$$\hat{q}(\mathbf{y}) = \begin{cases} 1; & C_{11}p(q=1|\mathbf{y}) + C_{01}p(q=0|\mathbf{y}) \\ & \quad < C_{10}p(q=1|\mathbf{y}) + C_{00}p(q=0|\mathbf{y}) \\ 0; & \text{otherwise,} \end{cases} \tag{D.22}$$

where $C_{qq'}$ is the cost resulting from the MSE of the estimator (D.21) with LOS decision q' applied under the true LOS condition q.

Implementation of the approximate MMSE estimators requires, in contrast to the approximate MAP estimators, evaluation of certain integrals. In case no closed-form expression can be obtained, this can be done fairly accurately by using standard numerical integration methods. We remark that the cost functions in (D.22) can be computed using Monte Carlo methods and stored for each considered parameter setting of the power delay profile. Therefore, range estimators with LOS detection require additional computational effort and storage compared to the estimators without LOS detection in (D.19) and (D.20).

4 Numerical Performance Evaluation

The invoked approximations of the likelihood function naturally impair the estimation performance. It is, however, unclear which of the estimators suffers the most. Note that the theoretical result that the MMSE estimator achieves lower MSE than all other estimators, e.g. the MAP estimator, does not hold for the approximate MMSE estimators. Thus, we rely on Monte Carlo simulations for assessing which of the above estimator yields the lowest MSE.

We compare the performance of the proposed estimators in terms of root-mean-squared-error (RMSE) and probability of LOS detection error. In addition, we compare them to "genie-aided" estimators obtained from (D.16) and (D.21) by inserting the true q value for $\hat{q}$. The genie-aided estimators provide lower bounds on the RMSE. As a study case, we simulate an OFDM communication system operating in the channel defined in the next subsection. Table D.1 reports the parameter settings used for the simulations.

Table D.1: Simulation Settings

OFDM system:
Bandwidth: 9 MHz, $N = 100$,
$\Delta f = 15\,\text{kHz}$, $T_{\text{cp}} = 5.4\,\mu\text{s}$, SNR $= \frac{E[\lvert a_i\rvert^2]}{\sigma^2}$,
Equal power and equal spacing pilot signal is used.
Channel parameters:
Homogenous Poisson point process: $\rho(\tau) = \rho_0$, $\lambda = 360\,\text{ns}$,
$\kappa = 2$; Average no. of paths: $\mu_L = \rho_0 T_{\text{cp}}$, $r \sim \mathcal{U}[0, 100]\text{m}$.
Results obtained from 3000 Monte Carlo trials are displayed.

4.1 Simulation Scenarios and Related Analytical Results

To reflect the situation where the user terminal (to be localized) can appear at any distance within an interval, we assume that the prior of range r is uniform on $[0, r_{\max}]$. Inspired by Turin's channel model, we assume that the random excess delays form a Poisson point process. For simplicity, we assume that the process is homogeneous, i.e. $\rho(\tau) = \rho_0$. The conditional second moments of the path gain are modeled as

$$\sigma_\alpha^2(\tau) = \begin{cases} C\kappa; & \tau = 0 \\ C\exp(-\frac{\tau}{\lambda}); & 0 < \tau < T_{\text{cp}} \\ 0; & \text{otherwise} \end{cases},$$

where parameter κ determines the power of the LOS component, λ denotes the root-mean-square (RMS) delay spread of the multi-path term, and C is selected to normalize the channel power gain. The cyclic prefix length, T_{cp}, is assumed to be long enough such that the power of the path components with excess delays larger than T_{cp} becomes negligible. Accordingly, the delay power spectrum reads

$$P(\tau) = C\exp(-\frac{\tau}{\lambda})(\rho_0 \mathbb{1}(0 < \tau < T_{\text{cp}}) + \kappa p_{\text{LOS}}\delta(\tau)) \tag{D.23}$$

with $\mathbb{1}$ denoting indicator function and the conditional delay power spectrum is given by

$$P(\tau|q) = \begin{cases} C\exp(-\frac{\tau}{\lambda})(\rho_0 \mathbb{1}(0 < \tau < T_{\text{cp}}) + \kappa\delta(\tau)); & q = 1 \\ C\exp(-\frac{\tau}{\lambda})\rho_0 \mathbb{1}(0 < \tau < T_{\text{cp}}); & q = 0. \end{cases} \tag{D.24}$$

For the simulation, it is necessary to compute the covariance matrices $\mathbf{C}_{\boldsymbol{\xi}}$ and $\mathbf{C}_{\boldsymbol{\xi}|q}$:

$$[\mathbf{C}_{\boldsymbol{\xi}}]_{mn} = \kappa p_{\text{LOS}} C + \rho_0 g_{mn} \tag{D.25}$$

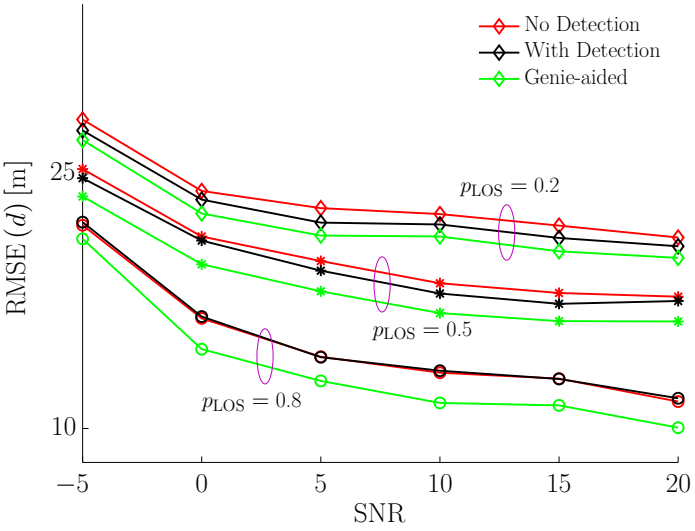


Fig. D.1: RMSE of the approximate MAP estimators versus SNR with p_{LOS} as a parameter: $\mu_L = 60$.

and

$$[\mathbf{C}_{\boldsymbol{\xi}|q}]_{mn} = \begin{cases} \kappa C + \rho_0 g_{mn}; & q = 1 \\ \rho_0 g_{mn}; & q = 0 \end{cases} \tag{D.26}$$

with

$$g_{mn} = C \frac{1 - e^{-(j2\pi(f_m - f_n) + \frac{1}{\lambda})T_{\text{cp}}}}{j2\pi(f_m - f_n) + \frac{1}{\lambda}}.$$

4.2 Evaluation of Ranging Accuracy

In the simulation, we obtain similar RMSEs for the approximate MAP estimators (D.14) and (D.15). The same observation holds for the approximate MMSE estimators (D.19) and (D.20). Therefore, we omit reporting the performance of (D.14) and (D.19).

Figs. D.1 and D.2 report the simulated RMSE versus SNR of the approximate MAP and MMSE estimators respectively for different values of p_{LOS}. It is apparent that the approximate MMSE estimators outperform the approximate MAP estimators. We observe that as p_{LOS} increases, the ranging accuracy improves.

To investigate the impact of the average number of path components on the estimators performance, we plot simulated RMSE versus μ_L in Figs. D.3

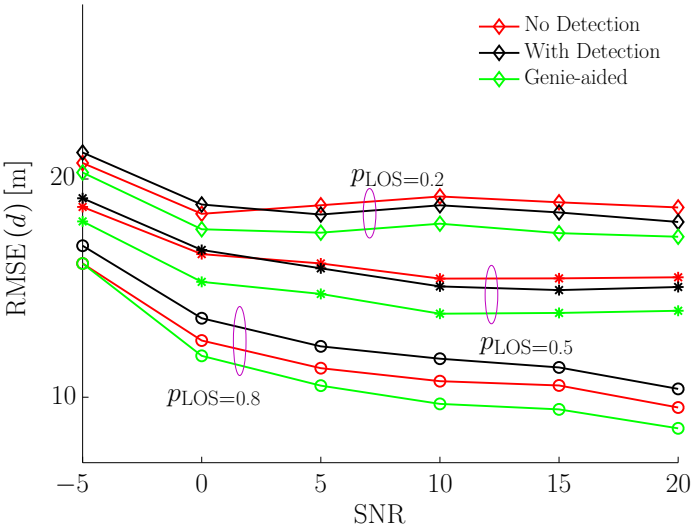


Fig. D.2: RMSE of the approximate MMSE estimators versus SNR with p_{LOS} as a parameter: $\mu_L = 60$.

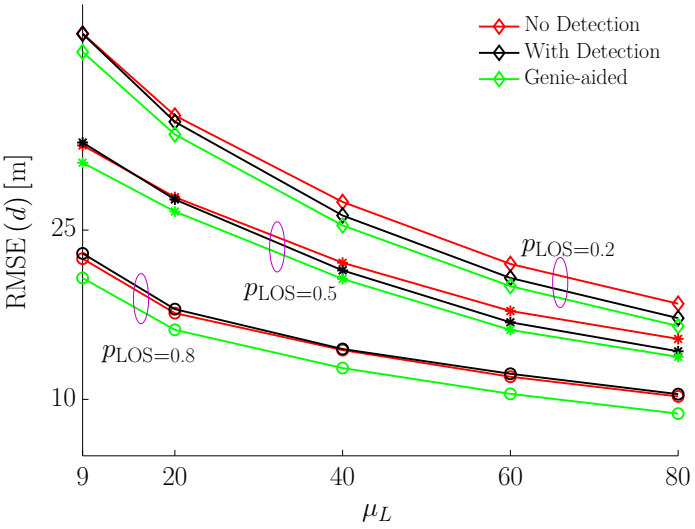


Fig. D.3: RMSE of the approximate MAP estimators versus the average number of paths with p_{LOS} as a parameter: SNR = 20 dB.

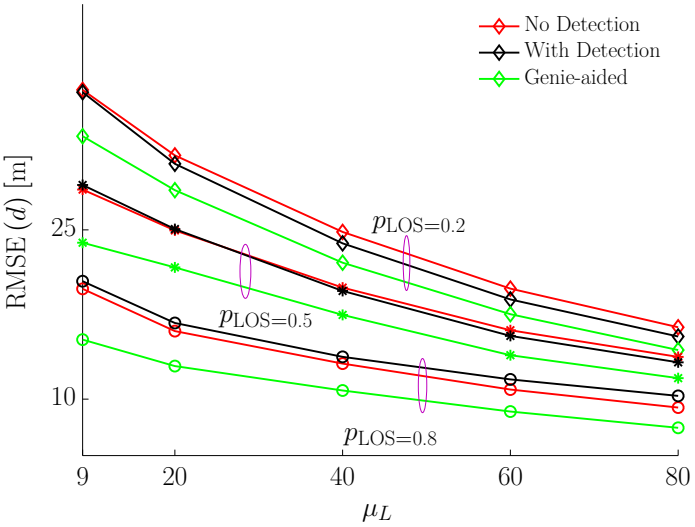


Fig. D.4: RMSE of the approximate MMSE estimators versus the average number of paths with p_{LOS} as a parameter: SNR = 20 dB.

and D.4, again when p_{LOS} is varied. We observe that, overall, the RMSE decreases with increasing μ_L. Similarly, the RMSE decreases as p_{LOS} increases as expected. As it is also observed in Figs. D.1 and D.2, the approximate MMSE estimators exhibit a higher ranging accuracy than the approximate MAP estimators.

Figs. D.1–D.4 indicate that including the LOS detector somewhat improves the ranging accuracy for low and medium values of p_{LOS}. However, for large p_{LOS}, the trend is different. For the approximate MAP estimator, there is no noticeable performance gain, while including the LOS detection in the approximate MMSE estimator degrades the performance. To investigate the cause of this behavior, we turn our attention to the performance of the detectors, see Fig. D.5. The probability of detection error seems rather high considering the prior information. This high value may be due to either the considered multipath channel or the pdf approximations applied to the design of the detectors.

Given these results, it seems obvious to ask whether or not the accuracy of the proposed methods can be improved by using better pdf approximations. Due to the fact that we cannot access the likelihood functions, lower bounds such as the Cramér-Rao bound, are not available. It is, therefore, unclear how much the estimation accuracy can be improved. To evaluate the importance of the impact of the approximations on the performance of the detectors, we applied them to signals generated according to their respective approximate

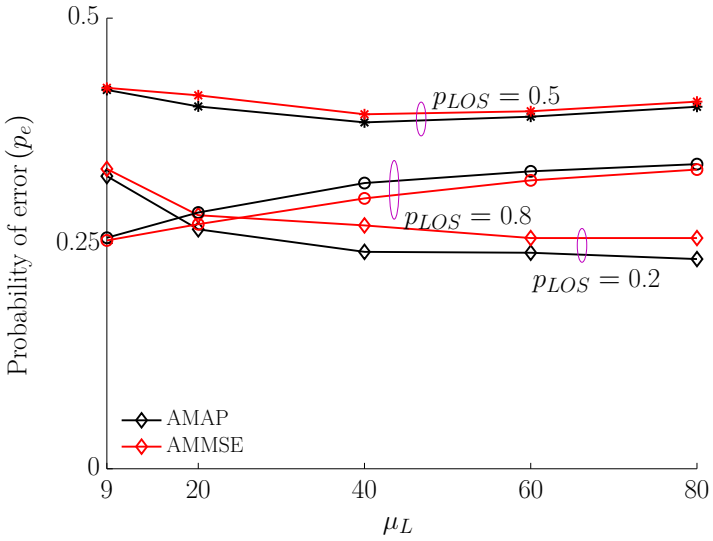


Fig. D.5: Probability of error versus the average number of paths at SNR = 20 dB.

pdfs. The results, not reported here, show error probabilities less than 7 % for all considered detectors with μ_L settings as given in Fig. D.5. We thus conclude that the accuracy of the pdf approximations indeed plays a major role. The potential performance gain in ranging accuracy obtained by better pdf approximations in the detector can be assessed by comparing the RMSE curves to those of the genie-aided methods as done in Figs. D.1–D.4. We conjecture that better pdf approximations can also increase the ranging accuracy of the estimators without detection.

5 Conclusion

We have proposed approximate MAP and MMSE estimators of the range with and without LOS detection. These estimators are derived by approximating the pdf of the received signal vector. The approximate MMSE estimators outperform the approximate MAP estimators in terms of RMSE. Using the proposed pdf approximations, we observe that including LOS detection in the estimators, while adding complexity, has no major impact on the ranging performance. Our simulation study indicates that there is a potential for improving the ranging performance by relying on better pdf approximations.

Acknowledgment

This work was supported in part by the EU FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306).

References

[1] A. Hero. Timing estimation for a filtered Poisson process in Gaussian noise. *IEEE Trans. Inf. Theory*, 37(1):92–106, Jan. 1991.

[2] B. Denis, J. Keignart, and N. Daniele. Impact of NLOS propagation upon ranging precision in UWB systems. In *IEEE Conference on Ultra Wideband Systems and Technologies*, pages 379–383, Nov. 2003.

[3] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win. Ranging with ultrawide bandwidth signals in multi-path environments. *Proc. IEEE*, 97(2):404–426, 2009.

[4] H. Wymeersch, S. Marano, W. Gifford, and M. Z. Win. A machine learning approach to ranging error mitigation for UWB localization. *IEEE Trans. Commun.*, 60(6):1719–1728, 2012.

[5] I. Guvenc and Chia-Chin Chong. A survey on TOA based wireless localization and NLOS mitigation techniques. *IEEE Communications Surveys Tutorials*, 11(3):107–124, Aug. 2009.

[6] J. Borras, P. Hatrack, and N. Mandayam. Decision theoretic framework for NLOS identification. In *Proc. IEEE Int. Technol. Conf.*, volume 2, pages 1583–1587 vol.2, May 1998.

[7] J. Khodjaev, Y. Park, and A. Saeed Malik. Survey of NLOS identification and error mitigation problems in UWB-based positioning algorithms for dense environments. *Annals of telecommunications*, 65(5-6):301–311, 2010.

[8] J. N. Kapur. *Maximum-entropy Models in Science and Engineering*. Wiley, 1989.

[9] J. Shen and A. F. Molisch. Indirect path detection based on wireless propagation measurements. *IEEE Trans. Wireless Commun.*, 11(12):4482–4493, Dec. 2012.

[10] Junyang Shen and A. F. Molisch. Indirect path detection based on wireless propagation measurements. *IEEE Trans. Wireless Commun.*, 11(12):4482–4493, Dec. 2012.

[11] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall PTR, 1998.

[12] L. Jing, T. Pedersen, and B. Fleury. Bayesian estimators for direct ranging in multi-path channels using OFDM pilot signals. 2014. Submitted to IEEE Trans. Wireless Commun.

[13] L. Jing, T. Pedersen, and B. Fleury. Direct ranging in multi-path channels using OFDM pilot signals. In *15th IEEE Int. Symposium on Signal Process. Advances in Wireless Commun.*, pages 150–154, June 2014.

[14] C. Mensing. *Location Determination in OFDM Based Mobile Radio Systems*. PhD thesis, Technische Universität München, 2013.

[15] R. Adam and P. A. Hoeher. Semi-blind channel estimation for joint communication and positioning. *10th Workshop on Positioning Navigation and Commun.*, pages 1–5, 2013.

[16] R. Zekavat and R. M. Buehrer. *Handbook of Position Location: Theory, Practice and Advances*. IEEE Series on Digital & Mobile Communication. Wiley, 2011.

[17] S. Marano, W. Gifford, H. Wymeersch, and M. Win. NLOS identification and mitigation for localization based on UWB experimental data. *IEEE J. Sel. Areas Commun.*, 28(7):1026–1035, 2010.

[18] A. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location. *IEEE Signal Process. Mag.*, 22(4):24–40, 2005.

[19] T. Wang, Y. Shen, S. Mazuelas, H. Shin and M. Z. Win. On OFDM ranging accuracy in multipath channels. *IEEE Systems Journal*, PP(99):1–11, 2013.

[20] D. Wang, M. Fattouche, and F. Ghannouchi. Fundamental limit of OFDM range estimation in a separable multipath environment. *Circuits, Systems, and Signal Processing*, 31(3), 2012.

[21] Y. Qi, K. Hisashi, and H. Suda. On time-of-arrival positioning in a multipath environment. *IEEE Trans. Veh. Technol.*, 55(5):1516–1526, 2006.

[22] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik. Non-line-of-sight identification and mitigation using received signal strength. *IEEE Trans. Wireless Commun.*, (99):1–1, Nov. 2014.